Summer 8-12-2013

Population Genetics and Phylogenetic Context of Weed Evolution in the Genus Amaranthus: Amaranthaceae)

Katherine Waselkov
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences

Evolution, Ecology and Population Biology

Dissertation Examination Committee:
Kenneth M. Olsen, Chair
James M. Cheverud
Allan Larson
Peter H. Raven
Barbara A. Schaal
Alan R. Templeton

Population Genetics and Phylogenetic Context of Weed Evolution in the Genus *Amaranthus*

(Amaranthaceae)

by

Katherine Elinor Waselkov

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2013

St. Louis, Missouri

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I gratefully acknowledge the financial support of the National Science Foundation (Graduate Research Fellowship), the National Geographic Society (Committee for Research and Exploration Grants 8901-11 and 9126-12), the Howard Hughes Medical Institute in cooperation with Washington University (HHMI Summer Undergraduate Research Fellowship to Nathan Regenold), and the Division of Biology and Biomedical Sciences (Evolution, Ecology, and Population Biology Program) at Washington University in St. Louis.

My committee members have provided assistance every step of the way. My co-advisor, Barbara Schaal, was particularly helpful during the early stages of my dissertation, when I wasn't really sure what I wanted to do. Barbara was always encouraging, whatever my project was at any given point, until I finally found my topic. Ken Olsen, my primary advisor, is largely responsible for pulling me through the end stages of my dissertation. He was always available to talk about analyses, writing, migraines, and proper care of the lab plants. Ken is also a great editor; almost daily, during the writing stages, he helped me to shorten grant applications, publications, and dissertation chapters and prompted me for more explicit ideas and less hyperbole. Alan Templeton, an excellent committee chair, cheered me up at the end of every committee meeting. Alan and Jim Cheverud provided statistics help at several stages, and Jim could always be counted on to point out my conflations of hypothesis and prediction. Allan Larson was my best listener during committee meetings and suggested improvements to the phylogenetic portion of the dissertation. And Peter Raven, the busiest man I know, made it to almost all of my committee meetings and asked thought-provoking, big picture questions. Peter also tapped his extensive network of contacts around the world to help me obtain specimens for the *Amaranthus* phylogeny.

The Schaal and Olsen lab members have been an invaluable source of feedback at lab meetings and companionship in lab. The older grad students who were here when I arrived, including Jason Londo, Brad Oberle, James Beck, and Taina Price, were founts of good advice and lab expertise, as were Templeton lab members Jennifer Neuwald, Michele Johnson, and Nick Griffin, and Josh Reece in the Larson lab. My graduate cohort and members of the next year's cohort, Genevieve Croft, Amber Burgett, Shih-Chung Hsu, Nic Kooyers, Pu Huang, Steve Kroiss, Melissa Simon, Amy Conley, and Katie Williams, offered their support, advice, and commiseration during my entire grad school career. I feel lucky to have interacted with all of them. In particular, Genevieve Croft was my partner in crime for most of grad school and made life in St. Louis much more interesting. Nic Kooyers and I jumped over a lot of the same hurdles together (classes, greenhouse work, and field work). He kept me grounded when I was in danger of going off the rails about one thing or another. I also received valuable mentorship from several postdocs: Eleanor Pardini, Briana Gross, Jared Strasburg, and Tong Pusadee all forgave my youth and were always available to help with lab work or life. Later in my dissertation research, Cindy Vigueira, Sara Wright, Claudia Henriquez, Jennifer Gruhn, and Jordan Teisher formed the "Scholsen" lab group I depended on to read my paper drafts.

Many people helped with my common garden experiments. The Knight lab at Washington University gave me great advice on ecological experiments. At Tyson Research Center, Kevin Smith, Pete Jamerson, Tim Derton, and Steve Bone provided space and assistance with tilling and fencing my plots. At Miami University's Ecology Research Center, Ann Rypstra and Rodney Kolb performed the same functions, and Rodney also saved me a ton of work by planting soybeans. Dave Smith at Miami University helped beyond end by taking data when I couldn't be in Ohio; that project never would have worked without him. At Tyson, Washington

University undergrads Nathan Regenold and Amal Al-Lozi woke up early, worked in the heat all day, and then did the same thing the next day for the entire summer of 2010. High school students Drew Wetherington and Asha Clark helped with the 2011 Tyson plot, and Susan Flowers and SIFT and TERF high school volunteers collected data and took down plots in 2010 and 2011.

I also received a lot of assistance with plant collecting and lab work. Rachel Kempen put up with my stressed-out behavior during several week-long road trips in which we had roughly one disaster per day, including briefly losing the car keys, losing track of each other in the woods, getting shot at, camping in some rough spots, and staying in some really sketchy hotels. Nathan Regenold had a similar experience for a shorter period of time during our collecting trip in Ohio and Indiana, with the addition of terrible allergies to our target plants and being chased by an angry dog. Nathan and Amal Al-Lozi helped with lab and greenhouse work, as did Drew Wetherington and Asha Clark. I also mentored Washington University undergrads David Vera-Vazquez (collecting around St. Louis and population genetics lab work), Alexis Boleda (lab work for the phylogeny project), and Sharon Jia (most of the PCR restriction enzyme screening for the herbicide resistance project). Ken Olsen's lab technician, Linda Small, has been endlessly helpful in lab. Besides training me in all the basic lab techniques, she was always there for troubleshooting. Carlo Lapid in the Templeton Lab helped me with the finicky computer program SSR Finder to design microsatellite primers.

In the Washington University Biology Department, Jane Ruiz and Paula Kiefel were always ready to work with my schedule and help me figure out what went wrong with my runs in the McDonnell Sequencing Center. Mike Dyer, Jeanne Cablish, and Darlene Branson in the Plant Growth Facility took very good care of my (occasionally enormous) plants and

accommodated any crazy project I came up with. In the accounting department, Melissa Evers patiently worked with the scanty receipt documentation from my collecting trips. Elsewhere in the department, Judy Musick, Betty Daues, and Mike Malolepszy were a great support team.

At other institutions, I really enjoyed working with Patrick Tranel's lab at the University of Illinois-Urbana/Champaign. Pat, his student Sukvinder Singh, and his postdoc Jerron Liu were great collaborators on the herbicide resistance project and provided me with transcriptome data to mine for microsatellites. My work on *Amaranthus tuberculatus* would not have been possible without all the previous work done in the Tranel lab. I also enjoyed visiting and exchanging ideas with Amy Lawton-Rauh and her graduate student Kristin Beard at Clemson University. The Missouri Botanical Garden was a great place to get an outside perspective on ideas. Peter Stevens, Peter Hoch, Iván Jiménez, and the Biodiversity and Ethnobotany Journal Clubs were all sources of inspiration during my dissertation research. My undergraduate advisor, Walter Judd at the University of Florida, is responsible for most of my publication record; I am very grateful to him for giving me such strong basic training in systematics.

Many, many people helped obtain species samples for the *Amaranthus* phylogeny. First up are Jorge Crisci, Anabela Plos and Pablo Simon from the Universidad Nacional de La Plata; Ana and Pablo made a special trip to collect *Amaranthus* in four Argentinian provinces. Fernando Zuloaga at the Instituto de Botánica Darwinion in Argentina sent me material and great pictures of additional Argentinian species. Mélica Muñoz at Museo Nacional de Historia Natural de Chile sent me samples of the rare Chilean endemic *A. looseri*. Carmen Bonifaz at the Universidad de Guayaquil helped me during my visit to the Guayaquil herbarium. Andrew Mitchell, John Hosking, and Barbara Briggs in Australia jointly helped me obtain *A. macrocarpus* material, and Jo Palmer at the Australian National Herbarium sent me herbarium

material of the remaining 10 Australian species. John Manning at the South African National Biodiversity Institute and Peter Goldblatt with the Missouri Botanical Garden helped obtain *A. praetermissus*; in Namibia, Remmie Moses and Steve Carr at the National Botanical Research Institute sent me *A. thunbergii* material. Warren Wagner at the U.S. National Herbarium sent me material of the possibly extinct Hawaiian endemic *A. brownii*. I will always regret that I couldn't extract it cleanly enough for good sequence data. Kurt Neubig sent me material of several *Amaranthus* species and outgroups from the University of Florida *Barcoding the Flora of Florida* database. David Brenner at the USDA GRIN database in Iowa was incredibly helpful and sent me material he wasn't releasing to the general public. The Mt. Cuba Center in Delaware sent me several live specimens of the endangered species *A. pumilus*. Mats Thulin at Uppsala University in Sweden sent me recently collected material of *Pleuropterantha revoilii*. Jim Solomon at the Missouri Botanical Garden herbarium allowed destructive sampling of several herbarium specimens as a last resort, including the outgroup *Chamissoa altissima*. Finally, Thomas Borsch and Uno Eliasson sent reprints of their work on *Amaranthus* and Amaranthaceae, and Néstor Bayón at the Universidad Nacional de La Plata very kindly allowed use of his unpublished revision of the monoecious *Amaranthus* species in my work. Many more people helped me contact the people listed, or kept an eye out for *Amaranthus* species during their own collecting trips.

During my Galápagos fieldwork, I was constantly bothering Roby Pépolas, Sonia Cisneros, and Solanda Rea at the Charles Darwin Foundation, but they always came through with permits and logistical assistance. Patricia Jaramillo and Frank Bungartz at the CDF helped me find an Ecuadorian assistant both years and helped me navigate the CDF herbarium. Washington Tapia and Galo Quezada at Galápagos National Park made it possible to obtain

permits for wherever I needed to go in the islands. Steve Blake, now at Tyson Research Center, aided with essentials like getting me on the right plane and buying groceries during my first trip. Ecuadorian undergraduate students Jorge Vera and Pedro Moreta were excellent field assistants, and Captain Lenín Cruz and the crew of *Pirata* helped me make the most of a constrained schedule and sketchy GPS coordinates on the second trip. Loren Sackett was my translator and photographer extraordinaire during the second trip and made the expedition a lot of fun.

Last but not least, I thank my parents Greg and Linda Waselkov, my little brothers Nicholas and Peter, my grandmother Elizabeth Burge, and my extended family in Missouri. I have been totally obsessed with plants for the last 28 years and you put up with me anyway, which I hope you'll keep doing.

ABSTRACT OF THE DISSERTATION

Population genetics and phylogenetic context of weed evolution in the genus *Amaranthus*

By Katherine Elinor Waselkov

Doctor of Philosophy in Evolution, Ecology, and Population Biology

Washington University in St. Louis, 2013

Professor Kenneth M. Olsen, Chairperson

Agricultural weeds have evolved to compete aggressively with domesticated plants in agricultural environments. Although the evolution of invasiveness has been studied extensively in natural ecosystems, few comparable studies have been conducted using agricultural weeds. In this dissertation, I used the genus *Amaranthus* to examine agricultural weed evolution over different evolutionary time scales, ranging from fitness measurements within a single species to a genus-wide, macroevolutionary analysis.

To explore the recent evolution of agricultural invasiveness, I studied a native Midwestern species, *A. tuberculatus* (waterhemp), which has become an aggressive agricultural weed only within the last several decades. I used microsatellite markers to investigate the present-day population structure of *A. tuberculatus*. To assess intraspecific variation in agricultural adaptation, I conducted a common garden study measuring the relative fitness of plants from across the species range in experimental soybean plots. I discovered two genetic subpopulations. The 20th century invasion of Midwestern agricultural fields was due to the eastward migration of the "western" genetic subpopulation, which has high competitive fitness in soybean fields and which may have been preadapted to the agricultural environment.

Waterhemp has rapidly evolved resistance to multiple classes of herbicides. The role of

native Midwestern riverbank populations in this process is unknown. I screened agricultural and riverbank populations of *A. tuberculatus* in Ohio for a common agricultural resistance mutation, using a combination of herbicide resistance phenotyping, PCR genotyping, and gene sequencing. I found that the most common agricultural mutation was indeed present in riverbank populations, suggesting that these native populations may serve as a reservoir of resistance alleles.

Finally, I constructed a phylogeny for the genus *Amaranthus* to investigate traits associated with the evolution of weediness. *Amaranthus* is a worldwide genus of 70 species, with no previous generic phylogeny. I included 58 species and two outgroups, sequenced at four nuclear genes and two chloroplast regions, in my molecular phylogenetic analyses using maximum parsimony, maximum likelihood, and Bayesian inference. Weediness exhibits no phylogenetic signal in *Amaranthus*; however, using non-phylogenetic statistical tests, I found associations of weediness with several morphological and ecological traits in the genus.

INTRODUCTION OF THE DISSERTATION

A major question in evolutionary biology is how introduced species invade natural habitats in their new range. This focus on natural habitats is well deserved, as invasive species are a primary concern in conservation biology; they are ranked as the second-most important threat to biodiversity worldwide, and billions of dollars are spent in the U.S. alone each year on invasive control efforts (Pimentel et al., 2000; Pimentel et al., 2005). Research is beginning to reveal the importance of evolutionary processes in the success or failure of invasions in natural systems (Lee, 2002; Parker et al., 2003; Bossdorf et al., 2005; Lee et al., 2007). And yet, some of the most dramatic instances of invasive species evolution have occurred in agricultural ecosystems (Baker and Stebbins, 1965; Clements et al., 2004; Neve et al., 2009; Vigueira et al., 2013). Agricultural weed evolution is of great interest to farmers and scientists alike. Innovations like herbicide resistance and improved competitive ability can cost many millions of dollars in crop yield losses every year (Jordan and Jannink, 1997; Neve et al., 2009). Furthermore, selection in these agricultural habitats is often extremely strong, causing rapid weed evolution (Baker, 1974; Tranel and Wright, 2002; Weinig, 2005). Nonetheless, the potential of agricultural weed systems to answer evolutionary questions about invasiveness remains largely untapped.

Despite a large amount of research on invasive species in many taxonomic groups, some fundamental evolutionary questions remain unanswered, including: What traits pre-adapt a species to become invasive? What is the relative importance of adaptation vs. plasticity for establishment in the new range? And what role does genetic variation play in invasion success? (Facon et al., 2006; Richards et al., 2006; Van Kleunen and Johnson, 2007; Dlugosch and Parker, 2008; Schlaepfer et al., 2010). Part of the reason for a lack of consensus may be that a multitude of community interactions occur with the introduced species, making invasions of natural

ecosystems complex and varied, so that generalizations are difficult (Facon et al., 2006). Simpler, human-managed ecosystems, such as crop fields, present alternative starting points for evolutionary invasive research because they have only one primary community member, and they experience strong selective forces that stem from known human activities (Clements et al., 2004; Smith et al., 2006). Some agricultural weed species evolved in concert with agriculture, and thus have coexisted with humans for centuries (Harlan, 1965). However, just as for invasives of natural habitats, increased dispersal and changes in land use lead to new opportunities for the evolution of "weediness" (Sauer, 1972; Warwick, 1990; Fuhrer, 2003).

Agriculture in the U.S. has changed drastically since the turn of the 19th century. Weedy plant species have always been associated with agriculture, distinguished from domesticated plants primarily by their lesser usefulness and lesser dependence on humans for survival (Det Wet and Harlan, 1975; Ghersa et al., 1994). Prior to agricultural intensification in the U.S., a diversity of weed species existed in any one crop field. Mechanization of nearly every step of farming and technological advances in soil improvement during the 20th century allowed a vast expansion in the amount of land under cultivation and homogenization of agricultural ecosystems (Ghersa et al., 1994). Furthermore, the detrimental impact of weed competition on crop yields became more important as profit margins in farming grew ever slimmer (due to the enormous inputs of fertilizer and herbicide required to maintain modern agroecosystems), leading to a proliferation of chemical classes of herbicides to obtain total control of weeds, and strongly selecting for weed species that could rapidly evolve herbicide resistance (Maxwell et al., 1990; Ghersa and Martinez-Ghersa, 1991). Late in the 20th century, the implementation of widespread no-till or conservation tillage agriculture, as well as the introduction of herbicide resistant crops, led to another shift in the weed species present in fields (Swanton et al., 1993;

Owen, 2008). Consequently, there are a number of plant species that have only recently become problematic in agricultural environments (Maillet and Lopez-Garcia, 2000).

Most studies to date that have investigated the origin and evolution of agricultural weeds have focused either on weeds that are related to the crop (which have the potential to obtain crop genes — including transgenes — via introgression) or on the acquisition of herbicide resistance by agricultural weed species and the molecular basis of this dramatic adaptation (for reviews, see Ellstrand et al., 1999; Jansieniuk et al., 1996). Weeds with domesticated relatives can arise through hybridization between domesticates and sympatric wild species, or through the evolution of a feral form of the domesticate ("de-domestication"), and weeds that originated through these mechanisms can be found in several systems (e.g., Burger et al., 2006; Olsen et al., 2007; Fénart et al., 2008). Agricultural weeds that have obtained genes from sympatric crop relatives are also well-known in several groups, including canola, radishes, and the sorghum/Johnson grass complex (Warwick et al., 2003; Morrell et al., 2005; Campbell et al., 2006). The evolutionary genetics of herbicide resistance has been the subject of several theoretical papers (e.g., Maxwell et al., 1990; Jasieniuk et al., 1996), which have largely focused on the ways in which farm management can slow down resistance evolution. Weed scientists are also intensely interested in discovering the genes and genetic changes involved in herbicide resistance for the same reason, and also to aid in the design of new herbicides (e.g., Patzoldt et al., 2006; Gaines et al., 2010).

Few studies have focused on tracing the origin of an agricultural weed without sympatric domesticated relatives, derived from a wild species which was until recently found only in natural habitats (but see Menchari et al., 2007; Kane and Rieseberg, 2008). This latter type of study is important both as a model for understanding invasion in general, and for creating guidelines for agricultural researchers attempting to decrease the likelihood of new aggressive

weeds arising. A relatively new weed invasion would still have the genetic signatures of the invasion event(s) in extant populations (e.g. Dlugosch and Parker, 2008). Ideally, the source populations for the new weed species should be known, as well as whether crop-weed hybridization was a potential avenue to permit adaptation (Schierenbeck and Ellstrand, 2009). Furthermore, a phylogenetic approach is seldom applied to researching the evolution of "weediness," despite its proven usefulness in understanding both traits correlated with invasion success and the potential interaction of the invasive with other community members (e.g. Burns, 2004; Parker and Gilbert, 2004).

My model agricultural weed species, *Amaranthus tuberculatus* (Moq.) Sauer (Amaranthaceae), or waterhemp, fits the necessary criteria for this type of study. It belongs to a genus that contains several widespread, agriculturally problematic weeds, as well as non-weedy species, which can provide a phylogenetic framework to examine the evolution of traits involved in weediness over a broad time scale. Furthermore, waterhemp is the bane of Midwestern corn and soybean farmers, and is the leading cause of crop yield loss in their fields (Steckel, 2007), and yet until recently it was not found in agricultural environments (Sauer, 1957). Waterhemp is native to riverbanks in the Midwestern United States, and was first noticed as an invasive in crop fields in the 1950s. Range expansion and/or hybridization between populations from different areas of the species' range are implicated in this invasion (Sauer, 1957; 1972). Waterhemp has become especially problematic in recent years due to the evolution of widespread herbicide resistance, and it is unknown whether most forms of resistance have a fitness cost in the absence of herbicide application (Tranel and Trucco, 2009). Riverbank populations in close proximity with agricultural populations may therefore harbor alleles for herbicide resistance, and could recolonize fields where resistance has been eradicated. This recent invasion allows me to

examine what roles population history, hybridization, and adaptation play in the evolution of invasiveness.

Study System

The Genus *Amaranthus*

Amaranthus is a genus of ~70 species in the Amaranthaceae, with species native to every continent; its greatest diversity is in warm temperate, subtropical, and tropical regions (Mosyakin and Robertson, 2003). It contains widespread weeds, restricted endemics, endangered species, and domesticated species (Sauer, 1950). General features of the genus include an annual life history; herbaceous habit; reduced unisexual flowers with male and female flowers on the same plant (monoecy) or different plants (dioecy); wind pollination; and tiny seeds that are typically dispersed by wind, water, or birds (Mosyakin and Robertson, 2003). In the most recent taxonomic work, the genus is divided into three subgenera: *Acnida* (all dioecious species), *Amaranthus,* and *Albersia* (Mosyakin and Robertson, 1996).

The genus has garnered interest in the past mainly for its domesticated species and its agricultural weed species. *Amaranthus caudatus* (domesticated in the Andes), *A. cruentus* (domesticated in Guatemala), and *A. hypochondriacus* (domesticated in central Mexico) have been the subject of many studies aiming to resolve the question of their phylogenetic origin, as well as many studies interested in crop potential and improvement. These grain amaranths were important in the Aztec and Incan empires, but were suppressed by the Spanish during colonial times (Sauer, 1950). Their nutritional and agricultural properties were not rediscovered until the 1970s: researchers found an almost complete complement of amino acids in the grain, as well as relatively high disease resistance and some drought resistance (Grubben and van Sloten, 1981).

Two semi-domesticated *Amaranthus* species are also eaten as vegetables in Europe, Asia, and Africa: *A. tricolor* (Asian origin) and *A. blitum* (European origin) (Mosyakin and Robertson, 2003).

No clearly-resolved phylogeny of the entire genus *Amaranthus* has been published. Previous phylogenetic work in *Amaranthus* has either involved only a subset of species, or produced very low-resolution results (e.g. Lanoue et al., 1996; Xu and Sun, 2001). Experiments in hybridization between weedy *Amaranthus* species have determined that *A. palmeri* and *A. tuberculatus* (both dioecious species) yield practically no fertile offspring when crossed, whereas *A. hybridus* (a monoecious species) and *A. tuberculatus* produce some fertile F1 individuals (Murray, 1940; Trucco et al., 2007; Trucco et al., 2009). These results suggest that the dioecious weed species may not be each other's closest relatives, and that deciphering the phylogenetic relationships between *Amaranthus* species could generate new hypotheses about the potential for gene flow between agricultural weed species.

Nine *Amaranthus* species are listed as "invasive or noxious weeds" in the USDA Plants Database, and an additional 20 species are listed as "agricultural weeds" in the Global Compendium of Weeds (USDA, NRCS, 2010; Randall, 2007). It is unknown whether morphological and physiological traits associated with invasion of agricultural ecosystems evolved once or several times within the clade, due to the lack of a generic phylogeny. Testing for phylogenetic signal in the evolution of agricultural weediness in *Amaranthus* can reveal whether it is necessary to control for shared evolutionary history when pinpointing adaptive "weedy" traits. Many important morphological and ecological traits are included in species descriptions from two major literature sources (Mosyakin and Robertson, 2003; Bayón, in review), and in particular, data associated with plant size, breeding system, seed dispersal,

geographic range, and habitat selection are available, which correspond to traits involved in Baker's "ideal weed" characteristics (Baker, 1974) and to traits examined in similar studies of invasive species of natural ecosystems (e.g., Jenkins and Keller, 2011). In testing for the association of these traits with agricultural invasiveness, several different metrics of "weediness" (again, based on the literature) should be used, to account for variation in the degree of agricultural invasiveness in the species different authors call weeds.

Waterhemp, *Amaranthus tuberculatus*

My focal species for the population-level studies of this genus is *Amaranthus tuberculatus*, an annual, wind-pollinated dioecious species (Mosyakin and Robertson, 1996). It is native to the Midwestern U.S. and its natural habitat includes muddy margins of rivers, streams and lakes, but it has adapted readily to man-made disturbed areas such as roadsides. It has very few competitors in these habitats (Mosyakin and Robertson, 2003). *Amaranthus tuberculatus* was first recorded as an agricultural weed in Illinois cornfields in the early 1950s (Sauer, 1957) and has become a weed of major concern since the 1990s, when herbicide resistance was first discovered (Trucco et al., 2009). Today, the species has evolved resistance to four herbicide types: acetolactate synthase (ALS)-inhibitors, protophyrinogen oxidase (protox)-inhibitors, photosystem II (PSII) -inhibitors, p-hydroxyphenylpyruvate dioxygenase (HPPD) inhibitors, and glyphosate (Patzoldt et al., 2005; Legleiter and Bradley, 2008; Hausman et al, 2011).

Historically, Sauer treated *A. tuberculatus* as occurring east of the Mississippi River and considered populations west of the Mississippi a different species, *A. rudis*. He distinguished them by morphological characters, the clearest of which was dehiscence or indehiscence of the fruit. Sauer hypothesized that the northeastward expansion of *A. rudis* during the 1940s and

1950s and subsequent introgression with *A. tuberculatus* in the area just east of the Mississippi led to the development of agroecotypes, or "weedy" forms of *A. tuberculatus,* in the 1950s (Sauer, 1957). His morphological characters showed intrapopulation variation only in the putative hybrid zone. Today, these weedy agroecotypes are most problematic in Missouri and Illinois corn and soybean fields, while only native riverbank populations are found in eastern Ohio, Michigan, and most of Ontario (Costea et al., 2005). Some later authors have treated these taxa as one highly variable species, *A. tuberculatus* (Pratt and Clark, 2001), while others treat them as varieties (var. *rudis* and var. *tuberculatus*), recognizing that the variation is geographically structured (Costea and Tardif, 2003). I will treat them as varieties for this dissertation.

Resistance to ALS-inhibiting herbicides has become extremely widespread in Midwestern *A. tuberculatus* since the 1990s, making waterhemp the most problematic weed of Illinois corn and soybean fields (Trucco et al., 2009). Patrick Tranel's lab at the University of Illinois-Urbana-Champaign has recently shown that some weed populations of waterhemp have evolved resistance to multiple herbicide types, sometimes showing multiple mutations at the same herbicide target site (Tranel et al., 2004; Patzoldt et al., 2005). Often only one or a few base pair changes cause resistance to a particular herbicide. ALS resistance appears to have no fitness cost in herbicide-free environments in some *Amaranthus* species (Sibony and Rubin, 2002; but see Tardif et al., 2006), although fitness tests have not been conducted with resistant *A. tuberculatus*. The Tranel lab is also sequencing the transcriptome of *A. tuberculatus*, which increases the genetic tools available for this system (Lee et al., 2009; Riggins et al., 2010).

Dissertation Overview

This dissertation explores the evolution of agricultural invasiveness at several evolutionary time scales. In Chapter 1, to place agricultural weed evolution in a phylogenetic context, I reconstruct the phylogeny of *Amaranthus* to examine the evolution of traits associated with agricultural invasiveness. In Chapter 2, I use *Amaranthus tuberculatus* to investigate the population structure of agricultural invasion, to test hypotheses about the importance of hybridization and range expansion in invasive evolution in a recently arisen agricultural weed. In Chapter 3, I determine whether genetic adaptation to crop field environments (other than herbicide resistance) has taken place in this invasive weed since it arose. Finally, in Chapter 4, I test the hypothesis that waterhemp populations in natural environments contain agriculturally-adaptive herbicide-resistance alleles. Chapter 2 has been written as a manuscript and is currently in review at the American Journal of Botany.

LITERATURE CITED

BAKER, H. G. 1974. The evolution of weeds. *Annual Review of Ecology and Systematics* 5: 1-24.

BAKER, H. G., AND G. L. STEBBINS, [eds.] 1965. The genetics of colonizing species. Academic Press, New York, NY, USA.

BAYÓN, N. D. In review. Revisión taxonómica de las especies monoicas de *Amaranthus* L. (Amaranthaceae): Amaranthus subg. Albersia y Amaranthus subg. Amaranthus. *Annals of the Missouri Botanical Garden*. Used with permission.

BOSSDORF, O., H. AUGE, L. LAFUMA, W. E. ROGERS, E. SIEMANN, AND D. PRATI. 2005. Phenotypic and genetic differentiation between native and introduced plant populations. *Oecologia* 144: 1-11.

BURGER, J. C., S. LEE, AND N. C. ELLSTRAND. 2006. Origin and genetic structure of feral rye in the western United States. *Molecular Ecology* 15: 2527-2539.

BURNS, J. H. 2004. A comparison of invasive and non-invasive dayflowers (Commelinaceae)

across experimental nutrient and water gradients. *Diversity and Distributions* 10: 387–397.

CAMPBELL, L. G., A. A. SNOW, AND C. E. RIDLEY. 2006. Weed evolution after crop gene introgression: greater survival and fecundity of hybrids in a new environment. *Ecology Letters* 9: 1198-1209.

CLEMENTS, D. R., A. DITOMMASO, N. JORDAN, B. D. BOOTH, J. CARDINA, D. DOOHAN, C. L. MOHLER, S. D. MURPHY, AND C. J. SWANTON. 2004. Adaptability of plants invading North American cropland. *Agriculture, Ecosystems and Environment* 104: 379-398.

COSTEA, M., S. E. WEAVER, AND F. J. TARDIF. 2005. The biology of invasive alien plants in Canada. 3. *Amaranthus tuberculatus* (Moq.) Sauer var. *rudis* (Sauer) Costea & Tardif. *Canadian Journal of Plant Science* 85: 507-522.

COSTEA, M., AND F. J. TARDIF. 2003. Conspectus and notes on the genus *Amaranthus* in Canada. Rhodora 105: 260-281.

DE WET, J. M. J., AND J. R. HARLAN. 1975. Weeds and domesticates: evolution in the man-made habitat. *Economic Botany* 29: 99-107.

DLUGOSCH, K. M., AND I. M. PARKER. 2008. Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology* 17: 43l-449.

ELLSTRAND, N. C., H. C. PRENTICE, AND J. F. HANCOCK. 1999. Gene flow and introgression from domesticated plants into their wild relatives. *Annual Review of Ecology and Systematics* 30: 539-563.

FACON, B., B. J. GENTON, J. SHYKOFF, P. JARNE, A. ESTOUP, AND P. DAVID. 2006. A general eco-evolutionary framework for understanding bioinvasions. *Trends in Ecology and Evolution* 21: 130-135.

FÉNART, S., J.-F. ARNAUD, I. DE CAUWER, AND J. CUGUEN. 2008. Nuclear and cytoplasmic genetic diversity in weed beet and sugar beet accessions compared to wild relatives: new insights into the genetic relationships within the *Beta vulgaris* complex species. *Theoretical and Applied Genetics* 116: 1063-1077.

FUHRER, J. 2003. Agroecosystem responses to combinations of elevated CO2, ozone, and global climate change. *Agriculture, Ecosystems and Environment* 97: 1-20.

GAINES, T. A., W. ZHANG, D. WANG, B. BUKUN, S. T. CHISHOLM, D. L. SHANER, S. J. NISSEN, ET AL. 2010. Gene amplification confers glyphosate resistance in *Amaranthus palmeri*. *Proceedings of the National Academy of Sciences* 107: 1029-1034.

GHERSA, C. M., AND M. A. MARTINEZ-GHERSA. 1991. Cambios ecológicos en los agrosistemas

de la Pampa Ondulada: efectos de la introduccion de la soja. *Ciencia e Investigación* 2: 182-188.

GHERSA, C. M., M. L. ROUSH, S. R. RADOSEVICH, AND S. M. CORDRAY. 1994. Coevolution of agroecosystems and weed management. *BioScience* 44: 85-94.

GRUBBEN, G. J. H., AND D. H. VAN SLOTEN. 1981. Genetic resources of amaranths: a global plan of action. International Board for Plant Genetic Resources, IBPGR Secretariat, Rome, Italy.

HARLAN, J. R. 1965. The possible role of weed races in the evolution of cultivated plants. *Euphytica* 14: 173-176.

HAUSMAN, N. E., S. SINGH, P. J. TRANEL, D. E. RIECHERS, S. S. KAUNDUN, N. D. POLGE, D. A. THOMAS, AND A. G. HAGER. 2011. Resistance to HPPD-inhibiting herbicides in a population of waterhemp (*Amaranthus tuberculatus*) from Illinois, United States. *Pest Management Science* 67: 258-261.

JASIENIUK, M., A. L. BRÛLÉ-BABEL, AND I. N. MORRISON. 1996. The evolution and genetics of herbicide resistance in weeds. *Weed Science* 44: 176-193.

JENKINS, C., AND S. R. KELLER. 2011. A phylogenetic comparative study of preadaptation for invasiveness in the genus *Silene* (Caryophyllaceae). *Biological Invasions* 13: 1471-1486.

JORDAN, N. R., AND J. L. JANNINK. 1997. Assessing the practical importance of weed evolution: a research agenda. *Weed Research* 37: 237-246.

KANE, N. C., AND L. H. RIESEBERG. 2008. Genetics and evolution of weedy *Helianthus annuus* populations: adaptation of an agricultural weed. *Molecular Ecology* 17: 384-394.

LANOUE, K. Z., P. G. WOLF, S. BROWNING, AND E. E. HOOD. 1996. Phylogenetic analysis of restriction-site variation in wild and cultivated *Amaranthus* species (Amaranthaceae). *Theoretical and Applied Genetics* 93: 722–732.

LEE, C. E. 2002. Evolutionary genetics of invasive species. *Trends in Ecology and Evolution* 17: 386-391.

LEE, C. E., J. L. REMFERT, AND Y.-M. CHANG. 2007. Response to selection and evolvability of invasive populations. *Genetics* 129: 179-192.

LEE, R. M., J. THIMMAPURAM, K. A. THINGLUM, G. GONG, A. G. HERNANDEZ, C. L. WRIGHT, R. W. KIM, M. A. MICKEL, AND P. J. TRANEL. 2009. Sampling the waterhemp (*Amaranthus tuberculatus*) genome using pyrosequencing technology. *Weed Science* 57: 463-469.

LEGLEITER, T. R., AND K. W. BRADLEY. 2008. Glyphosate and multiple herbicide resistance in common waterhemp (*Amaranthus rudis*) populations from Missouri. *Weed Science* 56: 582-587.

MAILLET, J., AND C. LOPEZ-GARCIA. 2000. What criteria are relevant for predicting the invasive capacity of a new agricultural weed? The case of invasive American species in France. *Weed Research* 40: 11-26.

MAXWELL, B. D., M. L. ROUSH, AND S. R. RADOSEVICH. 1990. Predicting the evolution and dynamics of herbicide resistance in weed populations. *Weed Technology* 4: 2-13.

MENCHARI, Y., C. DÉLYE, AND V. LE CORRE. 2007. Genetic variation and population structure in black-grass (*Alopecurus myosuroides* Huds.), a successful, herbicide-resistant, annual grass weed of winter cereal fields. *Molecular Ecology* 16: 3161-3172.

MORRELL, P. L., T. D. WILLIAMS-COPLIN, A. L. LATTU, J. E. BOWERS, J. M. CHANDLER, AND A. H. PATERSON. 2005. Crop-to-weed introgression has impacted allelic composition of johnsongrass populations with and without recent exposure to cultivated sorghum. *Molecular Ecology* 14: 2143-2154.

MOSYAKIN, S., AND K. R. ROBERTSON. 1996. New infrageneric taxa and combinations in *Amaranthus* (Amaranthaceae). *Annals of Botanici Fennici* 33: 275-281.

MOSYAKIN, S., AND K. R. ROBERTSON. 2003. *Amaranthus*. *In* Flora of North America Editorial Committee [eds.], Flora of North America North of Mexico, Vol. 4. Magnoliophyta: Caryophyllidae, Pt. 1. Oxford University Press, New York, USA.

MURRAY, M.J. 1940. The genetics of sex determination in the family Amaranthaceae. *Genetics* 25: 409-431.

NEVE, P., M. VILA-AIUB, AND F. ROUX. 2009. Evolutionary-thinking in agricultural weed management. *New Phytologist* 184: 783-793.

OLSEN, K. M., A. L. CAICEDO, AND Y. JIA. 2007. Evolutionary genomics of weedy rice in the USA. *Journal of Integrative Plant Biology* 49: 811-816.

OWEN, M. D. K. 2008. Weed species shifts in glyphosate-resistant crops. *Pest Management Science* 64: 377-387.

PARKER, I. M, AND G. S. GILBERT. 2004. The evolutionary ecology of novel plant-pathogen interactions. *Annual Review of Ecology, Evolution and Systematics* 35: 675-700.

PARKER, I. M., J. RODRIGUEZ, AND M. E. LOIK. 2003. An evolutionary approach to understanding the biology of invasions: local adaptation and general-purpose genotypes in the weed *Verbascum thapsus*. *Conservation Biology* 17: 59-72.

PATZOLDT, W. L, P. J. TRANEL, AND A. G. HAGER. 2005. A waterhemp (*Amaranthus tuberculatus*) biotype with multiple resistance across three herbicide sites of action. *Weed Science* 53: 30-36.

PATZOLDT, W. L., A. G. HAGER, J. S. MCCORMICK, AND P. J. TRANEL. 2006. A codon deletion confers resistance to herbicides inhibiting protoporphyrinogen oxidase. *Proceedings of the National Academy of Sciences* 103: 12329-12334.

PIMENTEL, D., R. ZUNIGA, AND D. MORRISON. 2005. Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics* 52: 273-288.

PIMENTEL, D., L. LACH, R. ZUNIGA, AND D. MORRISON. 2000. Environmental and economic costs of nonindigenous species in the United States. *BioScience* 50: 53-65.

PRATT, D. B, AND L. G. CLARK. 2001. *Amaranthus rudis* and *A. tuberculatus*, one species or two? *Journal of the Torrey Botanical Society* 128: 282-296.

RANDALL, R. P. 2007. The Global Compendium of Weeds (http://www.hear.org/gcw/). AgWest and the Hawaiian Ecosystems at Risk (HEAR) project.

RICHARDSON, C. L., O. BOSSDORF, N. Z. MUTH, J. GUREVITCH, AND M. PIGLIUCCI. 2006. Jack of all trades, master of some? On the role of phenotypic plasticity in plant invasions. *Ecology Letters* 9: 981-993.

RIGGINS, C. W., Y. PENG, C. N. STEWART JR., AND P. J. TRANEL. 2010. Characterization of de novo transcriptome for waterhemp (*Amaranthus tuberculatus*) using GS-FLX 454 pyrosequencing and its application for studies of herbicide target-site genes. *Pest Management Science* 66: 1042-1052.

SAUER, J. D. 1950. The grain amaranths: a survey of their history and classification. *Annals of the Missouri Botanical Garden* 37: 561-631.

SAUER, J. D. 1957. Recent migration and evolution of the dioecious amaranths. *Evolution* 11: 11-31.

SAUER, J. D. 1972. The dioecious amaranths: a new species name and major range extensions. *Madroño* 21: 426-434.

SCHIERENBECK, K. A., AND N. C. ELLSTRAND. 2009. Hybridization and the evolution of invasiveness in plants and other organisms. *Biological Invasions* 11: 1093-1105.

SCHLAEPFER, D. R., M. GLATTLI, M. FISCHER, AND M. VAN KLEUNEN. 2010. A multi-species experiment in their native range indicates pre-adaptation of invasive alien plant species. *New Phytologist* 185: 1087-1099.

SIBONY, M., AND B. RUBIN. 2002. The ecological fitness of ALS-resistant *Amaranthus retroflexus* and multiple-resistant *Amaranthus blitoides*. *Weed Research* 43: 40-47.

SMITH, R. G., B. D. MAXWELL, F. D. MENALLED, AND L. J. REW. 2006. Lessons from agriculture may improve the management of invasive plants in wildland systems. *Fronteirs in Ecology and the Environment* 4: 428-434.

STECKEL, L.E. 2007. The dioecious *Amaranthus* spp.: here to stay. *Weed Technology* 21: 567-570.

SWANTON, C. J., D. R. CLEMENTS, AND D. A. DERKSEN. 1993. Weed succession under conservation tillage: a hierarchical framework for research and management. *Weed Technology* 7: 286-297.

TARDIF, F. J., I. RAJCAN, AND M. COSTEA. 2006. A mutation in the herbicide target site acetohydroxyacid synthase produces morphological and structural alterations and reduces fitness in Amaranthus powellii. *New Phytologist* 169: 251-264.

TRANEL, P. J, AND F. TRUCCO. 2009. 21st century weed science: a call for *Amaranthus* genomics. *In* C. N. Stewart, Jr. [ed.], Weedy and Invasive Plant Genomics. Blackwell Publishing, Ames, Iowa, USA, pp. 53-81.

TRANEL, P. J., W. JIANG, W. L. PATZOLDT, AND T. R. WRIGHT. 2004. Intraspecific variability of the acetolactate synthase gene. *Weed Science* 52: 236-241.

TRANEL, P. J., AND T. R. WRIGHT. 2002. Resistance of weeds to ALS-inhibiting herbicides: what have we learned? *Weed Science* 50: 700-712.

TRUCCO, F., D. ZHENG, A. J. WOODYARD, J. R. WALTER, T. C. TATUM, A. L. RAYBURN, AND P. J. TRANEL. 2007. Nonhybrid progeny from crosses of dioecious amaranths: implications for gene-flow research. *Weed Science* 55: 119-122.

TRUCCO, F., T. TATUM, A. L. RAYBURN, AND P. J. TRANEL. 2009. Out of the swamp: unidirectional hybridization with weedy species may explain the prevalence of *Amaranthus tuberculatus* as a weed. *New Phytologist* 184: 819-827.

USDA, NRCS. 2010. The PLANTS Database (http://plants.usda.gov, 17 November 2010). National Plant Data Center, Baton Rouge, Louisiana, USA.

VAN KLEUNEN, M., AND S. D. JOHNSON. 2007. South African Iridaceae with rapid and profuse seedling emergence are more likely to become naturalized in other regions. *Journal of Ecology* 95: 674-681.

VIGUEIRA, C. C., K. M. OLSEN, AND A. L. CAICEDO. 2013. The red queen in the corn: agricultural weeds as models of rapid adaptive evolution. *Heredity* 110: 303-311.

WARWICK, S. I. 1990. Allozyme and life history variation in five northwardly colonizing North American weed species. *Plant Systematics and Evolution* 169: 41-54.

WARWICK, S. I., M.-J. SIMARD, A. LÉGÈRE, H. J. BECKIE, L. BRAUN, B. ZHU, P. MASON, ET AL. 2003. Hybridization between transgenic *Brassica napus* L. and its wild relatives: *Brassica rapa* L., *Raphanus raphanistrum* L., *Sinapis arvensis* L., and *Erucastrum gallicum* (Willd.) O. E. Schulz. *Theoretical and Applied Genetics* 107: 528-539.

WEINIG, C. 2005. Rapid evolutionary responses to selection in heterogeneous environments among agricultural and nonagricultural weeds. *International Journal of Plant Science* 166: 641-647.

XU, F., AND M. SUN. 2001. Comparative analysis of phylogenetic relationships of grain amaranths and their wild relatives (*Amaranthus*; Amaranthaceae) using internal transcribed spacer, amplified fragment length polymorphism, and double-primer fluorescent intersimple sequence repeat markers. *Molecular Phylogenetics and Evolution* 21: 372-387.

CHAPTER 1

Molecular Phylogeny of *Amaranthus* (Amaranthaceae) and

Trait Associations with Weediness in the Genus

INTRODUCTION

The plant genus *Amaranthus* (Amaranthaceae) includes 60-70 species, with the bulk of the species (40-50) native to the Americas, but a handful native to Eurasia, South Africa, and Australia (Bayón, in review). The genus is characterized by the following traits: an annual or (rarely) short-lived perennial life history; alternate leaves; imperfect flowers (plants monoecious or dioecious) in compound dichasia packed into inflorescences; inflorescences terminal and/or axillary; generally three to five tepals and stamens; utricle or pyxidium fruit; and a base chromosome number of 16 or 17 (Mosyakin and Robertson, 2003). In addition, the genus has C_4 photosynthesis, unlike its closest related genera (Sage et al., 2007). Several species of *Amaranthus* are economically important: three species are domesticated pseudocereals popular in South America and South Asia, two species are grown as vegetable crops in Asia, and around eight species are problematic agricultural weeds worldwide (Mosyakin and Robertson, 2003). Müller and Borsch (2005) and Sage et al. (2007) place the genus in the Amaranthaceae, subfamily Amaranthoideae, tribe Amarantheae, subtribe Amaranthinae, closely related to the genera *Pleuropterantha* and *Chamissoa*. There is currently no well-supported, well-sampled phylogeny of the genus, despite its wide geographical distribution and close association with human activities.

The genus *Amaranthus* was first established by Linnaeus in 1753. Various parts of the genus were at one time recognized as separate genera, particularly the dioecious species and the monoecious species with dehiscent or indehiscent fruits (Linnaeus, 1753; Kunth, 1838). These genera were later placed within *Amaranthus* by Grenier and Godron (1856), Sauer (1955), and Robinson (1981), and are presently recognized as subgenera in the group by most authorities; *Amaranthus* includes subgenus Acnida, subgenus Amaranthus, and subgenus Albersia

(Mosyakin and Robertson, 1996; Costea et al., 2001). Subgenus Acnida is generally delimited to include all of the dioecious species of *Amaranthus*, whereas subgenus Amaranthus and subgenus Albersia split the monoecious species using a combination of inflorescence position, number of tepals, and fruit dehiscence (Mosyakin and Robertson, 1996; Bayón, in review). Several authorities have suspected that this infrageneric taxonomy may not correspond well to evolutionary history (Eliasson, 1988; Mosyakin and Robertson, 2003).

Previous phylogenetic work in the genus *Amaranthus* has either involved a very restricted sample of species, or produced very low-resolution results. In 1996, a study using 30 species available from the USDA was conducted using restriction-site analysis of three PCR-amplified loci (1 nuclear and 2 chloroplast) (Lanoue et al., 1996). Although 38 restriction endonucleases were used, only 14% of restriction site data were polymorphic, and the resulting phylogenetic trees had many polytomies. However, Lanoue et al.'s results did show several conserved groups in all trees, which placed species of the subgenus Acnida into two separate clades. Therefore, the best-sampled phylogenetic work in the genus, while very low resolution, does not support the monophyly of the taxonomic subgenera. Many other studies have involved limited reconstruction of relationships between species in subgenus *Amaranthus* (especially the *A. hybridus* species complex) using a variety of molecular markers, including RAPDs and isozymes (Chan and Sun, 1997); low-C_OT DNA sequences (Sun et al., 1999); *ITS* DNA sequences, AFLPs and ISSRs (Xu and Sun, 2001); and microsatellites (Mallory et al., 2008). These studies support the origin of the domesticated grain amaranths (*A. hypochondriacus*, *A. cruentus*, and *A. caudatus*) from *A. hybridus*, although some find evidence for lesser contributions from other species (e.g. Xu and Sun, 2001). Another study used AFLPs to study relationships among eight U.S. agricultural weeds (Wassom and Tranel, 2005). Most species in the genus have never been

analyzed genetically or phylogenetically. The general lack of phylogenetic resolution in previous analyses suggests that *Amaranthus* could be a very recently radiated genus, which might predispose DNA sequence datasets to incomplete lineage sorting. This is especially true for nuclear genes, which have effective population sizes twice the size of the effective population size of chloroplast DNA (or four times the size for dioecious species) (Templeton 2006).

The unusual pan-global distribution leads to interesting biogeographical questions regarding the history of diversificiaton in *Amaranthus*. The geographical region of origin of the genus and the relationships of the Old World species to the New World species are both unknown. Long-distance dispersal between continents is almost definitely involved in the radiation of the genus, as it is less than 65 million years old (Kadereit et al., 2003). Furthermore, there are three to four *Amaranthus* species native or endemic to the Galápagos Islands. Morphological similarities between these species and various other species in the genus have been noted by previous authorities, but it is currently unknown whether the genus radiated in the islands after a single colonization event, or arrived in the Galápagos multiple times (Eliasson, 1985; 1987).

My particular interest in *Amaranthus* stems from the many agricultural weeds in the genus. I view my work on the phylogeny as an opportunity to look for phylogenetic signal in the evolution of weediness, and to analyze associations of morphological and ecological traits with agricultural invasiveness. Evolution of invasive plants of natural ecosystems is often studied in a phylogenetic context (e.g., Burns, 2004; Muth and Pigliucci, 2006; Van Kleunen and Johnson, 2007; Van Kleunen et al., 2008; Fenesi et al., 2011), but agricultural weed evolution has seldom been approached the same way (but see Daehler, 1998; Brändle et al., 2003; Lososová et al., 2008). A few studies have attempted to find traits associated with weediness by analysis of the

agricultural weeds of a particular flora (Perrins et al., 1992; Sutherland, 2004). Many others have measured traits in weed species that are putatively associated with weediness, but have not compared them to those of their non-weedy congeners (e.g., seed dormancy, reviewed in Benech-Arnold et al., 2000). Phylogenetically-controlled analyses can assure that conclusions about weed-specific adaptations are not confounded with evolutionary history (Felsenstein, 1985). In the specific case of weediness-associated traits, Baker (1974) put forward testable hypotheses about the traits of ideal weeds or "colonizing species," which have since been interpreted to include ruderal plants, agricultural weeds, and invasive species. These hypotheses have been tested to some extent in invasives of natural ecosystems (reviewed in Pysek and Richardson, 2007), but few studies have examined these traits in agricultural weeds (but see Chaney and Baucom, 2012). A comprehensive, phylogenetically-based survey of these traits in agricultural weeds and invasive species revealed marked differences between traits adaptive for the two types of invaders (Daehler, 1998), which is not entirely surprising, given that agricultural ecosystems bear little resemblance to most natural ecosystems, especially in the regularity and frequency of disturbance.

I set out to reconstruct the phylogeny of *Amaranthus* in order to answer questions about the phylogenetic placement of agricultural weeds, but also to answer generally interesting questions about biogeographic relationships in the genus and the monophyly of the subgenera. I also collected information on morphological and ecological traits from the taxonomic literature on the species to test for associations of these traits with agricultural invasiveness in *Amaranthus*.

MATERIALS AND METHODS

Taxon Sampling

The genus *Amaranthus* contains 65 species total according to Bayón (in review). Fifty-seven of these species are sampled here (we treat *A. quitensis* as a separate species from *A. hybridus*, unlike Bayón's treatment), plus a known interspecific hybrid, a putative new species and an unidentified, possibly new species. Multiple accessions (including subspecies) of each species were sampled when material was available, for a total of 102 specimens of *Amaranthus* included in the phylogeny. In addition, two outgroup species from closely-related genera were included: two specimens of *Chamissoa altissima*, a Neotropical clambering shrub, and one specimen of *Pleuropterantha revoilii*, a North African shrub. These genera are the closest relatives of *Amaranthus* based on Sage et al.'s 2007 *matK/trnK*-based phylogeny of Amaranthaceae, with *Pleuropterantha* being the sister taxon to *Amaranthus* and *Chamissoa* being the sister taxon to *Pleuropterantha* + *Amaranthus*.

Species and subspecies included in the phylogenetic reconstruction are listed in Table 1.1, with taxonomic authorities. Also listed is the classification of each species in two recent taxonomic treatments of the genus, the Flora of North America treatment (Mosyakin and Robertson, 1996; 2003) and the complete treatment of the monoecious species (Bayón, in review). Finally, the geographic origin of each species is listed by continent and by area within continent. This species list excludes the two undetermined species. The first of these is a specimen from Argentina called "mystery species" in my analyses because of its unclear affinity to any other species morphologically or molecularly. The second is a putative new species from the Galápagos Islands, called "new species," with close molecular and morphological affinity to the Caribbean species *A. crassipes*. The specimens included in the phylogenetic reconstruction

are listed in Table 1.2 with their abbreviations in the phylogenetic trees (sometimes these abbreviations were further truncated as an artifact of the phylogenetic inference software). The source of each specimen is also provided in Table 1.2; many specimens were obtained from the USDA Germplasm Resources Information Network (GRIN) database, which has an extensive collection of wild and cultivated *Amaranthus* species. When herbarium material or specially-collected material was used, the collector, collection number, herbarium, and herbarium accession number are listed when available. The geographical provenance of each accession is also listed.

DNA Extraction and Sequencing

DNA was extracted from each sample with Qiagen DNeasy Plant Mini Kits (Qiagen Inc., Valencia, California, USA), except for the herbarium specimens. These were ground with liquid nitrogen, and then processed using a modification of Doyle and Doyle's (1990) CTAB plant extraction protocol. After the choroform extraction step, reagents and columns from the Invitrogen PureLink PCR Purification Kit (Life Technologies, Carlsbad, California, USA) were used to clean the DNA. Five volumes of Invitrogen binding buffer were added to the aqueous phase and mixed. Then this mixture was loaded onto the columns provided in the kit, and the columns were washed and eluted with Invitrogen wash buffer and then elution buffer. This procedure produced higher-quality, cleaner *Amaranthus* DNA from well-preserved herbarium material than did the Qiagen Plant Mini Kit.

Four nuclear genes and two chloroplast regions were amplified and sequenced for each specimen. The nuclear genes were *A36* (a predicted DEAD-box ATP-dependent RNA helicase), *G3PDH* (glyceraldehyde 3-phosphate dehydrogenase), *ITS* (internal transcribed spacers 1 and 2

and the intervening 5.8S ribosomal gene), and *Waxy* (granule-bound starch synthase). The chloroplast regions were *matK/trnK* (the maturase K gene and surrounding *trnK* intron) and *trnC-trnD* (intergenic region). Primers and read length of the aligned genes are listed in Table 1.3. *A36*, *ITS,* and *matK* primers were obtained from Amy Lawton-Rauh's lab at Clemson University. The *G3PDH* primers were redesigned after amplification with primers from Strand et al. (1997) to amplify one specific gene copy of the two *G3PDH* copies detected by cloning. The internal primers for *A36* and *G3PDH* were used only if the DNA quality was too poor to obtain a high-quality sequence read from the external primers. In contrast, the *matK/trnK* and *Waxy* internal primers were used for all species, due to the length of the amplified region.

PCR was performed on ABI GeneAmp 9700 thermocyclers (Applied Biosystems, Carlsbad, California, USA), in 25 uL reactions containing: 1X GoTaq Flexi Buffer (Promega, Madison, Wisconsin, USA), 2 mM MgCl2, 0.2 mM each dNTPs, 0.8 uM each forward and reverse primers, 0.125 uL GoTaq, 7.125 uL nanowater, 6.25 uL betaine, and from 2-4 uL genomic DNA. Amplification conditions were: 94ºC for 5 minutes, then 35 cycles of 94 ºC (30 seconds) denaturation, 50 ºC (30 seconds) annealing, 68 ºC (2 minutes) extension, and 72 ºC (7 minutes) final extension. PCR cleanup was performed with Invitrogen PureLink Quick PCR Purification Kits, according to the manufacturer's instructions but starting with 20-25 uL PCR products. Direct sequencing was performed in 12 uL reactions containing: 0.625X sequencing buffer, 0.27 uM primer, 1.0 uL PCR product, 1.0 uL BigDye version 3.0 terminator (Applied Biosystems), and 6.9 uL nanowater. Sequencing reaction conditions were: 96ºC for 1 minute, then 50 cycles of 96 ºC (10 seconds) denaturation, 50 ºC (5 seconds) annealing, and 60 ºC (4 minutes) extension. Sequences were cleaned with Sephadex columns (GE Healthcare, Piscataway, NJ, USA) and sequenced on the ABI Prism 3130x Genetic Analyzer (Applied

Biosystems). Cloning was performed for the dioecious species (which are obligately outcrossing, unlike the monoecious species, which are highly selfing (Murray, 1940)) and for *A. dubius*, the lone allotetraploid species. Sequencing of cloned PCR products was similar to direct sequencing, except with the intermediate steps of transformation of ligated PCR products into Z competent *E. coli* cells (Zymo Research Co., Irvine, California, USA), followed by plating and colony PCR. I obtained at least eight clones per species to distinguish and phase the two alleles for heterozygotes, and to eliminate SNPs and haplotypes resulting from PCR recombination or other replication error during cloning.

All sequences were combined into contigs, and quality scores were assigned with the "phred and phrap" function of BioLign 4.0.6.2 (Hall, 2005). If after several sequencing attempts, the quality of a particular base call was still ambiguous, this site was removed from the dataset. After automatic alignment in BioLign, sequence alignments were proofread by eye and edited if necessary. Gaps and indels were coded as missing data in all subsequent analyses, because of some uncertainty in homology of particular gaps (especially in *G3PDH*).

Phylogenetic Analyses

Phylogenetic trees were constructed using single genes and also the concatenated nuclear gene dataset and the concatenated chloroplast region dataset. For individual nuclear gene analyses, multiple alleles (if present) were included for the dioecious species and *A. dubius*, to detect incomplete lineage sorting. For concatenation of nuclear genes, multiple alleles for a single gene and single specimen were combined into a consensus sequence using IUPAC ambiguity codes for heterozygous sites. Because of the low phylogenetic informativeness of individual chloroplast genes and the complete linkage disequilibrium across the chloroplast

genome, *trnC-trnD* and *matK/trnK* were always analyzed as a concatenated unit. The chloroplast and nuclear datasets support different phylogenetic positions for many species in the genus (see Results); therefore, an analysis of all genes concatenated together was deemed inappropriate. Three methods were used to reconstruct trees: maximum parsimony (MP), maximum likelihood (ML), and Bayesian inference (BI).

Maximum Parsimony Analyses

Because of the computational intensity and duration of the parsimony analyses, the program PAUPRat (Parsimony ratchet searches using PAUP*: Sikes and Lewis, 2001) was implemented on the CIPRES Science Gateway platform (UC-San Diego, www.phylo.org) to search for the shortest tree for only the concatenated nuclear and concatenated chloroplast datasets. The ratchet parameters were set to 200 ratchets with 10 ratchet iterations per replicate, with 20 percent of characters perturbed each iteration, and a uniform weighting mode. The parsimony search parameters were set to a tree bisection-reconnection branch swapping algorithm, and the specimen "Chamissoa" was specified as the outgroup (because only one outgroup could be specified). The shortest trees were saved and converted to a Newick format. Then the Phylip program Consense (Felsenstein, 2005) was run to produce a 50% majority-rule consensus tree from the shortest trees.

PAUP* v 4.0 (Swofford, 1998) was used to produce parsimony bootstrap trees for the concatenated datasets using the "faststep" search command, and also using the "heuristic" search command while saving 100 trees per pseudoreplicate, due to the extreme duration of heuristic searches with an unrestricted maximum number of trees. The results of these two methods were compared and found to be very similar (with differences in bootstrap support); therefore, only

the results of the "100 max trees" method, which gave slightly higher bootstrap support for most partitions, are presented. A 50% majority-rule consensus tree was created from 100 bootstrap pseudoreplicates in PAUP*. Trees were visualized and manipulated to appear similar (by branch rotation) for easier comparison using MEGA 5.2.1 (Tamura et al., 2011).

Maximum Likelihood Analyses

I ran maximum likelihood analyses for single gene and concatenated datasets. For these, I used RAxML 7.3.1-HPC BlackBox (Stamatakis, 2006) on CIPRES. For RAxML on this platform, the molecular model is fixed by the program to GTRCAT for bootstrapping, and I chose GTR + Γ for the final tree inference for all data sets (as suggested by the programmer). I allowed RAxML to halt bootstrapping automatically, and specified all specimens of *Chamissoa* and *Pleuropterantha* as outgroups. The tree with the highest maximum likelihood was saved with branch lengths estimated. A 50% majority-rule consensus tree was created from the bootstrap pseudoreplicate trees using Consense on CIPRES. Trees were visualized and manipulated to appear similar in MEGA.

Bayesian Inference Analyses

I ran Bayesian analyses for single gene and concatenated datasets. For these, I used the program MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003) on CIPRES. For this program on CIPRES, the molecular model can be changed to fit the data by the user: accordingly, the program MrModeltest 2.3 (Nylander, 2004) was run in PAUP* for each data set (individual genes and concatenated datasets) and the model with the highest AIC (Akaike Information Criterion) and hierarchical likelihood ratio test (hLRT) values was selected. If these measures

each ranked a different model highest, separate MrBayes analyses were run using each model and the results compared. For the concatenated nuclear dataset, a partitioned analysis (specifying a different molecular model for each gene in the dataset) was run.

In addition to specifying the model of nucleotide evolution, I set the outgroup as "Chamissoa," and used uniform priors for the analyses. Each analysis consisted of two independent runs of four chains each (three heated, one cold) with a chain temp of 0.2, with 1,000,000 generations to start (first 25% of values discarded as burnin), sampling the Markov chain every 1000 generations. If the standard deviation of split frequencies was not ~0.01, or if the potential scale reduction factor was not ~1.0 for all parameters, after a million generations, then the analysis was run longer. The output from MrBayes was a 50% majority-rule consensus tree including branch lengths and posterior probability values for each partition in the tree. Trees were visualized and manipulated to appear similar in FigTree 1.4.0 (Rambaut, 2012). In addition, I used FigTree to show the phylogenetic placement of major clades, and to plot taxonomy (subgenus), biogeography (continent of origin), and agricultural weediness on the Bayesian trees.

Topology Testing

To test taxonomic hypotheses about incomplete lineage sorting and incongruence between trees, and to test the monophyly of the Galápagos Islands species and their closest relatives, I used Templeton's nonparametric test (1983) implemented in PAUP*. Heuristic searches consisting of 20 replicates (with a rearrangement limit of 50,000 per replicate) were performed in PAUP* to find the shortest unconstrained tree for each dataset. Then topological constraints were applied and the heuristic search was repeated. The shortest constrained tree was

compared to the shortest unconstrained tree using Templeton's test.

Trait Associations with Weediness

To test associations of morphological and ecological traits with agricultural invasiveness in the genus, I collected data on traits from the literature. I focused on morphological traits available from the literature associated with plant growth/size, breeding system, and seed dispersal, following Baker's (1974) list of traits that make an "ideal" weed. I found data on 12 morphological and ecological traits in two main literature sources (Mosyakin and Robertson, 2003; Bayón, in review) with supplementary sources for obscure species (Hunziker, 1951; Hunziker, 1965; Hunziker, 1966; Brenan, 1981; Palmer, 2009) (see Table 1.4). The quantitative traits are self-explanatory, except for the "number of GBIF cells occupied" trait. This is a proxy for the amount of geographical area a species occupies: specimen records with geographical coordinates are mapped in the Global Biodiversity Information Facility (GBIF) database (data.gbif.org), and I counted the number of 1º x 1º (latitude x longitude) cells occupied by the species according to these records. The derivation of qualitative traits from species descriptions was straightforward, with the exception of the following traits. Fruit dispersal by water was inferred indirectly from an inflated utricle or tepals persistent in fruit and fleshy/spongy/fused at the fruit base. I inferred that a species' range had been expanded by humans if it is listed as "introduced" in any region of the world in species' descriptions. Finally, the habitat traits were binary traits: for example, if a species was mentioned as occurring on beaches (no matter where else it occurs), it was scored as "beaches."

Agricultural invasiveness was scored in three different ways, again based on published species descriptions from the same sources. The agricultural weed rank metric had three levels:

the first level was for species that never occur in agricultural fields; the second for species that occasionally occur in agricultural fields and/or are opportunistic agricultural weeds; and the third level was for species that are mentioned as strongly associated with agriculture or problematic agriculturally. The second metric was "agricultural weed status," a binary character that divided species into non-weeds and weeds (species described as sometimes or frequently occurring in agricultural environments). The third metric was also binary, but "problematic weed status" divided species into problematic weeds (equal to the third level in the ranking metric) and non-problematic weeds or non-weeds.

I tested each of these weediness metrics for phylogenetic signal using the packages "ape," "picante" and "caper" in the programming language R (Paradis et al., 2004; Kembel et al., 2010; Orme et al., 2011). For each analysis, I loaded a tree based on the Bayesian concatenated nuclear gene tree that was pruned before phylogenetic analysis to match the taxonomic units for which I had agricultural weediness data. For the agricultural weed rank metric, I calculated Blomberg's K using "picante" and "ape," and I used a similar test, "phylo.d," in the "caper" package, to calculate the statistic D to detect phylogenetic signal in the binary traits agricultural weed status and problematic weed status. All tests showed that weeds are not distributed differently than expected by chance in the phylogeny, so I rejected the hypothesis that weediness in *Amaranthus* contains a phylogenetic signal (see Results). This meant that associations between weediness and other traits in the genus could be estimated with simple statistical tests, without phylogenetic independent contrasts.

To test for associations between agricultural invasiveness and qualitative traits, I used PASW Statistics 18.0.0 for Windows (SPSS Inc., Hong Kong, China). First I tested each qualitative trait for normality and log- or square root-transformed non-normal variables. I then

used independent-samples t-tests to detect differences in the means of traits in each agricultural weed status and problematic weed status category. ANOVAs were used to complete the equivalent tests for agricultural weed rank. To test for associations between agricultural invasiveness and quantitative traits, I performed chi-square tests with contingency tables in Microsoft Excel (Microsoft Co., Redmond, Washington, USA).

RESULTS

DNA Sequencing

Complete DNA sequences were successfully obtained for *A36* for 105 accessions (100% of all accessions), for *G3PDH* for 102 accessions (97%), for *ITS* for 105 accessions (100%), for *Waxy* for 96 accessions (91%), and for the chloroplast genes for 103 accessions (98%). Despite several attempts, one accession of *Chamissoa altissima* ("Chamissoa3") could not be sequenced for *G3PDH* or *Waxy*. In addition, the first 160 bp of *G3PDH* were too low-quality to align for the remaining accession of *C. altissima*, and a 100 bp intronic section of *G3PDH* in *Pleuropterantha revoilii* proved to be unalignable to *Amaranthus* and was removed. Only the last 420 bp of *Waxy* would align for the remaining *C. altissima* accession, and a 330 bp intronic section of *P. revoilii* was unalignable for *Waxy* and was removed. A high-quality sequence of *Waxy* could not be obtained for one accession of *Amaranthus clementii* ("ClemCran"), and this accession was removed from analysis. Several smaller portions of the same intronic section of *Waxy* (less than 100 bp) could not be sequenced for *A. cochleitepalus,* "DeflexC", "RhombeusG", *A. scariosus*, and "Undulat580". For one accession of *C. altissima* ("Chamissoa3") and for *A. urceolatus*, a 500 bp section in the *matK* region of *trnK/matK* could not be amplified successfully with the internal primers.

Phylogenetic Trees

Chamissoa altissima and *Pleuropterantha revoilii* were recovered as outgroups to *Amaranthus* without constraint for all individual-gene datasets and for the concatenated genes, with the exception of *G3PDH*, which nests *P. revoilii* within *Amaranthus* when *C. altissima* is set as the outgroup, and vice versa. *Amaranthus* thus had to be constrained as a monophyletic group in order to keep *Amaranthus* monophyletic for *G3PDH* trees. All unconstrained trees strongly support the monophyly of *Amaranthus*, and *C. altissima* and *P. revoilii* are on substantially longer branches than any *Amaranthus* species in RAxML and MrBayes analyses (see Figures 1.1 and 1.2).

Maximum Parsimony

The PAUPRat analysis of the concatenated nuclear dataset yielded 207 most-parsimonious trees of 1666 steps. The consistency index (CI) of each tree was 0.703 (CI excluding autapomorphies=0.615), and the retention index (RI) was 0.907. The 50% majority-rule consensus tree created by Consense is shown in Figure 1.3. The PAUPRat analysis of the concatenated chloroplast dataset recovered 209 most-parsimonious trees of 519 steps, with the CI = 0.871 (0.811 excluding autapomorphies), and the RI = 0.932. The 50% majority-rule consensus tree created by Consense is shown in Figure 1.4.

One hundred bootstrapped pseudoreplicates were sampled from the concatenated nuclear dataset and the concatenated chloroplast dataset, and used to create 50% majority-rule consensus trees in PAUP* (shown in Figure 1.5 and Figure 1.6 respectively). The bootstrapped consensus trees are generally less well-resolved, with the only differences in topology resulting from

bootstrap support values lower than 50% for some clades included in the consensus best tree.

Maximum Likelihood and Bayesian Inference

For the concatenated nuclear dataset, the highest-likelihood tree is very similar to the PAUPRat best consensus tree, with differences in the placement of the Galápagos clades (see below) (Figure 1.7). For the concatenated chloroplast dataset, the highest-likelihood tree is more dissimilar from the PAUPRat best consensus tree, with differences in the placement of several small clades and the arrangement of species within larger clades (Figure 1.8). Bootstrapped trees from PAUP* and RAxML have slight differences in bootstrap support yielding slight differences in topology for the nuclear dataset, and substantially better resolution of clades in RAxML than in PAUP* for the chloroplast dataset (Figures 1.9 and 1.10).

For all datasets, Bayesian trees are also very similar to bootstrapped trees from PAUP* and/or RAxML, with comparable support values to RAxML and slightly higher support values than PAUP* for some clades (where >85% is considered a high bootstrap value [70-84% is considered moderate support, Hillis and Bull, 1993] and 0.95-1.0 is considered a high posterior probability value [0.90-0.94 is considered moderate support]). Bayesian consensus trees for the nuclear and chloroplast concatenated datasets are shown in Figures 1.11 and 1.12.

Molecular models chosen for each species based on highest AIC and hLRT values are listed in Table 1.4. Two different models were chosen for the concatenated chloroplast dataset, and both models were used in separate MrBayes runs: the model GTR + Γ yielded substantially higher posterior probability values for several clades. For brevity, only posterior probability values from the Bayesian chloroplast tree based on the molecular model with the highest hLRT value are presented in the text.

RAxML and MrBayes trees based on single nuclear genes showed incomplete lineage sorting among alleles of the dioecious species and *A. dubius*, e.g., alleles from the same accession are more closely related to alleles of another accession or species, rather than being monophyletic. These relationships differed between genes (as expected), but were consistent between phylogenetic reconstruction methods and molecular models (see Figures 1.13-1.16).

Major Clades

Eurasian/South African/Australian (ESA)+South American Clade:

The ESA clade contains all of the Eurasian, South African, and Australian species in *Amaranthus*. It is supported by the concatenated nuclear and chloroplast datasets (posterior probability [PP] = 0.99 for both), and by the single nuclear genes *Waxy* (PP=1), *A36* (PP=0.98), and *G3PDH* with the inclusion of one accession of *A. blitoides* (see further discussion below) (0.99). This clade is subtended by 11 South American *Amaranthus* species. The inclusion of these species in a larger clade with the ESA clade is supported by the concatenated nuclear (PP=0.97) and chloroplast (PP=1) datasets, and by the single nuclear genes *G3PDH* (PP=1) and *A36* (PP=0.93), and weakly by *Waxy* (PP=0.88) (Figures 1.11-1.14 and 1.16).

Hybridus Clade:

The Hybridus clade consists of *A. hybridus* and its domesticated and wild or weedy relatives from the Americas. It is supported by the concatenated nuclear (PP=0.99) and chloroplast (PP=1) datasets, and by all single nuclear genes: *A36* (PP=0.99), *G3PDH* (PP=0.99), *Waxy* (PP=0.98), and more weakly by *ITS* (PP=0.79) (Figures 1.11-1.16).

Dioecious/Pumilus Clade(s):

All dioecious species of *Amaranthus* are included in this group, except for *A. palmeri* and

A. watsonii in the nuclear trees (see Hybridus clade discussion below). The group is supported as monophyletic by the concatenated nuclear (PP=1) dataset, but not by the chloroplast dataset, which divides the dioecious species and *A. pumilus* into two separate clades, and places *A. palmeri* and *A. watsonii* into one of them (Figures 1.11 and 1.12). A single Dioecious/Pumilus clade is supported by *ITS* (PP=0.99) and *A36* (PP=0.98), but not by *Waxy* (which splits the clade into two clades) or *G3PDH* (which splits the clade into three clades) (Figures 1.13-1.16).

Galápagos Clade(s):

The remaining species appear in various combinations in trees based on different genes. Only one gene, *G3PDH*, recovers these species as a monophyletic group (PP=0.99), which I call the Galápagos clade (because all Galápagos species occur in this clade) (Figure 1.14). The other datasets do not support this clade, but several analyses recover smaller "Galápagos clades" within it. The concatenated nuclear and chloroplast datasets and the ITS gene support a clade containing the Galápagos species *Amaranthus anderssonii,* the putative new Galápagos species, and three Caribbean species, henceforth referred to as the Anderssonii clade. The concatenated nuclear dataset and the ITS and Waxy genes support a clade containing the Galápagos endemic *A. sclerantoides* plus three western North American species, henceforth called the Sclerantoides clade. Finally, the concatenated nuclear and chloroplast datasets and the A36, ITS and Waxy genes support a clade including the Galápagos species *A. squamulatus* with a mainland South American species and a western North American species, henceforth called the Squamulatus clade (Figures 1.11, 1.12, 1.13, 1.15, and 1.16).

There is some evidence for a sister-group relationship between the Hybridus clade and the Dioecious/Pumilus clade, as shown in the concatenated nuclear tree (PP=0.98) (Figure 1.11). This relationship is supported by *A36* (PP=0.99) and more weakly by *ITS* (PP=0.83), as well as

by *Waxy* with the inclusion of the Sclerantoides and Squamulatus Galápagos clades (PP=0.92), but is not supported by *G3PDH* and the relationship between the clades is unresolved in the concatenated chloroplast tree (Figures 1.12-1.16).

The position of the other major clades of *Amaranthus* relative to each other is uncertain, especially the position of the Galápagos clade(s) relative to the ESA+South American clade and the Hybridus+Dioecious/Pumilus clade, and the position of these two latter clades relative to each other. A polytomy of these clades within a monophyletic *Amaranthus* appears to be the best-supported representation of evolutionary relationships based on this study. The major clades and areas of major disagreement between the nuclear and chloroplast datasets are shown in Figures 1.11 and 1.12.

Relationships within Major Clades

ESA+South American Clade:

There are very few well-supported relationships within this group. Species with multiple accessions are recovered as monophyletic (or unresolved) in the trees based on the concatenated nuclear and chloroplast datasets. Within the ESA clade, the two South African species are recovered as closely related to each other, but the concatenated datasets support different relationships between them: in the chloroplast tree, they are sister species, while in the nuclear tree, they are in a clade with *A. graecizans*, a European species. The remaining European and Australian species are largely unresolved: it appears that the taxon called "AffCuspid," which was identified as morphologically somewhat similar to *A. cuspidifolius*, may be *A. cochleitepalus* or an unrecognized closely-related species, supported by the concatenated nuclear dataset (PP=0.94), and the concatenated chloroplast dataset (PP=0.98). The nuclear tree places *A.*

macrocarpus, A. mitchelli, and *A. centralis* in a highly supported clade (PP=0.99), while the chloroplast places *A. centralis* with *A. cochleitepalus* instead (PP=0.97).

The gene *G3PDH* places one accession of *A. blitoides* into the ESA clade as the sister taxon to *A. graecizans* and the other accession with *A. albus* and its closest relatives. This is unexpected, given that the two accessions form a monophyletic species in the *A. albus* clade in every other tree except the chloroplast tree (where the species is unresolved). The *G3PDH* sequencing of *A. blitoides* was repeated several times to ensure that the species' names were not confused. Possibly a hybridization event between this particular accession of *A. blitoides* and *A. graecizans* occurred at the USDA GRIN database, but if this is the case, it is unclear why this anomalous relationship would not show up in other genes.

Regarding the basal South American species in this clade, the four species *A. crispus, A. persimilis, A. standleyanus*, and the mystery species from Argentina are a monophyletic group (PP=1 in the nuclear tree and the chloroplast tree). The identity of the latter species is still a mystery, as it is not very genetically similar to any of the other three species. The relationships among *A. deflexus, A. muricatus, A. viridis*, and *A. vulgatissimus* are puzzling: various pairs of the species are highly supported as sister taxa to each other by different genes. The inclusion of all four species in a monophyletic group is supported strongly by the chloroplast (PP=1), and weakly by the concatenated nuclear genes (PP=0.82). *Amaranthus kloosianus* and *A. looseri* are strongly supported as sister taxa (PP=1) and together as the sister taxon to the remainder of the ESA+South American clade (PP=0.98) by the chloroplast, but in the nuclear tree, *A. looseri* is the sister taxon to the remainder of the clade (PP=0.98), and *A. kloosianus* has an unresolved position in the clade.

Hybridus Clade:

Relationships within this clade are also poorly resolved. The clade includes *A. dubius*, a known allotetraploid that originated through hybridization between two species in this clade (Sauer, 1967). This hybrid origin is reflected in the disagreement between the concatenated chloroplast and nuclear datasets in the placement of the species: *A. dubius* is strongly supported as the sister species to *A. spinosus* by the chloroplast tree, while the two *A. dubius* accessions are not monophyletic in the nuclear tree, but appear in a clade with the core Hybridus group (discussed below). *G3PDH, A36*, and *Waxy* were cloned for *A. dubius*, and one allele of each accession is placed into the core Hybridus group, while the other is placed with *A. spinosus*, for each gene (Figures 1.13, 1.14, and 1.16). The exclusion of *A. dubius* does not change the topology or significantly change the posterior probabilites of the concatenated nuclear tree (data not shown).

Another major disagreement between the chloroplast and nuclear trees is unexpected: *A. palmeri* and *A. watsonii* (considered probable sister species based on morphology) appear as a sister clade to *A. spinosus* with strong support in the nuclear tree (PP=1) and are a part of the larger Hybridus clade (PP=0.99). This relationship is also supported by all four single nuclear genes (which exhibit incomplete lineage sorting for these species (Figures 1.13-1.16)). The chloroplast dataset, in contrast, places *A. palmeri* and *A. watsonii* in a clade with *A. pumilus* (PP=1) and with several of the other species in the Dioecious/Pumilus clade (*A. acanthochiton, A. tuberculatus, A. floridanus, A. arenicola*, PP=0.98).

The nuclear tree supports *A. spinosus* as the sister taxon to the remainder of the Hybridus clade, but the chloroplast tree instead places a clade consisting of *A. retroflexus, A. wrightii, A. powellii*, and *A. scariosus* as the sister taxon to the remaining Hybridus clade species. This same clade occurs as the next branching clade in the nuclear tree, except that it also includes *A.*

acutilobus. Within this clade, *A. retroflexus* and *A. wrightii* are supported as sister taxa by the chloroplast dataset (PP=1), and also by a single nuclear gene (*Waxy*, PP=0.91). Multiple accessions were sampled for *A. spinosus, A. powellii*, and *A. wrightii*, and each of these species is monophyletic or unresolved in every tree, with the caveat that the accession "PowelliZ" seems to be misidentified, and belongs in the core Hybridus group.

The core Hybridus group consists of *A. hybridus, A. hypochondriacus, A. caudatus, A. cruentus, A. quitensis*, and the accession "PowelliiZ". This group is weakly resolved in most trees, with or without the inclusion of *A. dubius, A. scariosus*, and *A. acutilobus*. Species and accessions in this group form a polytomy in most trees, with the exception of a few relationships. In the nuclear tree, the two accessions of *A. hypochondriacus* form a monophyletic group (and are strongly supported as the sister group to *A. hybridus*, PP=0.99), but this is not the case in the chloroplast tree, where one accession is placed with *A. spinosus*. The remaining accession of *A. hypochondriacus* is highly supported as belonging to another clade with *A. hybridus* and *A. cruentus* (PP=0.99). Also, the two accessions of *A. caudatus* are monophyletic in the nuclear tree (PP=0.99), but not in the chloroplast tree, where "CaudatusARG" is strongly supported as the sister lineage to *A. quitensis* (PP=0.99).

Dioecious/Pumilus Clade(s):

As mentioned above, this grouping of species appears as several phylogenetically disparate clades in the trees based on two single nuclear genes and the chloroplast tree. The *Waxy* gene splits the clade into two clades, one consisting of *A. australis, A. cannabinus*, and three *A. tuberculatus* alleles (PP=1), and the other consisting of *A. pumilus*, the remaining *A. tuberculatus* allele, and all other dioecious species (PP=1). The *G3PDH* gene splits the clade into three clades: one consists of *A. floridanus* and one *A. tuberculatus* allele (PP=1), the second

consists of *A. australis, A. cannabinus*, and *A. greggii* (PP=0.99), and the third consists of *A. pumilus*, the remaining *A. tuberculatus* alleles, and the rest of the dioecious species (PP=1). Finally, the concatenated chloroplast dataset supports two clades, one containing *A. australis* and *A. cannabinus* (PP=1) and the other containing *A. pumilus* and the rest of the dioecious species (PP=0.99).

The dioecious species were all cloned for each nuclear gene, and alleles of a single accession often exhibit incomplete lineage sorting (Figures 1.13-1.16). This obscures the relationships between species in this group, and leads to conflict between the nuclear and chloroplast trees, but some relationships between individual species/accessions are well-resolved. First, *A. pumilus*, a monoecious species, is undoubtedly closely related to the dioecious species, as its inclusion in a clade with some or all dioecious species is highly supported by every gene. The three accessions of *A. pumilus* are very similar genetically and form a clade in all trees. Furthermore, the sister-species relationship between *A. australis* and *A. cannabinus* is supported by both nuclear and chloroplast datasets (PP=1 for both trees). Finally, the two accessions of *A. greggii* form a monophyletic (or unresolved) group in all trees.

<u>*Galápagos Clade(s):*</u>

The species in this group are the native Galápagos species and their close relatives, all from the Americas. Only *G3PDH* places all of these species into a monophyletic group, but several smaller clades within the group are better supported by several genes. First, the Galápagos species *A. anderssonii* is very closely related to the Caribbean species *A. polygonoides*: in fact, the sequences are identical for several genes, and every gene and concatenated dataset places the two in a highly-supported monophyletic group. The putative new Galápagos species ("NewSps") is very closely related to the Caribbean species *A. crassipes*,

and is recovered as its sister species by two nuclear genes, the concatenated nuclear dataset, and the chloroplast dataset. These two small clades are placed together in a clade (the Anderssonii clade) along with the Mexican/Caribbean species *A. tamaulipensis* in the concatenated nuclear and chloroplast trees, albeit not with high support (PP=0.74 and 0.87, respectively).

The western U.S. species *A. albus* and *A. californicus* are closely related, and occur together in a highly-supported monophyletic group in every tree. In most trees, this clade also includes *A. blitoides* and *A. sclerantoides* (a Galápagos species), to form the Sclerantoides clade: these species form a separate clade from *A. albus* and *A. californicus* for the chloroplast dataset, and the *A36* nuclear gene includes *A. blitoides* in a clade with *A. albus* and *A. californicus* but excludes *A. sclerantoides* (PP=0.96). *G3PDH* only weakly supports the four species as a clade (PP=0.51), but this is probably due to the exclusion of one accession of *A. blitoides* placed in the ESA clade (see above).

Finally, the Galápagos species *A. squamulatus* is closely related to two other species, the southwestern U.S. species *A. fimbriatus* and the South American species *A. urceolatus*, which I call the Squamulatus clade. Somewhat surprisingly, *A. fimbriatus* and *A. urceolatus* appear to be each others' closest relatives (PP=1 in the nuclear and chloroplast trees), with *A. squamulatus* being the sister taxon to both of them combined (PP=1 in the nuclear tree, PP=0.96 in the chloroplast tree). The hybrid taxon *A. x tucsonensis* is not placed with high confidence in most trees, but one of its parents may be *A. fimbriatus*, as it is placed into this clade by the concatenated nuclear dataset (PP=0.99).

Topology Tests

I used Templeton's nonparametric test (1983) implemented in PAUP* to test several

hypotheses derived from the phylogenetic analyses. First, I wanted to test whether incomplete lineage sorting in the dioecious species was highly supported in the individual gene trees, by comparing the length of the shortest tree constrained to keep alleles from the same accessions monophyletic to the length of the shortest unconstrained tree for each gene. I also tried constraining the trees to keep each dioecious species monophyletic, which allowed alleles to be nonmonophyletic as long as the accessions for each species were.

Results of these tests depended on the gene. *A36* trees were not significantly longer with alleles or species constrained to be monophyletic (203 and 206 steps respectively, compared to 201 steps in the unconstrained tree), and neither were *ITS* trees (349 and 347 steps respectively, compared to 346 steps in the unconstrained tree). But *G3PDH* trees were significantly longer than the basic tree (480 steps, already constrained to keep the ingroup monophyletic) with alleles constrained as monophyletic (521 steps, $P<0.0001$) or species constrained as monophyletic (509 steps, $P=0.0001$). And *Waxy* trees were also significantly longer than the unconstrained tree (720 steps) with alleles constrained as monophyletic (770 steps, $P<0.0001$) or species constrained as monophyletic (757 steps, $P=0.0001$).

Next, I wanted to test whether the placement of *A. palmeri* and *A. watsonii* is highly supported as different by the nuclear dataset and the chloroplast dataset. When the nuclear tree is constrained to place *A. palmeri* and *A. watsonii* in a monophyletic group with *A. pumilus* (as in the chloroplast tree), the constrained tree is significantly longer (1722 vs. 1666 steps in the unconstrained tree, $P<0.0001$). When the nuclear tree is constrained to place the two species in the broader monophyletic group of *A. acanthochiton, A. arenicola, A. floridanus, A. greggii, A. tuberculatus,* and *A. pumilus*, the constrained tree is still significantly longer (1695 steps, $P=0.0001$). When the chloroplast tree is constrained to put the two species in a monophyletic

group with *A. spinosus* (as in the nuclear tree), the tree is significantly longer (533 vs. 519 steps in the unconstrained tree, P=0.0043). However, when the constrained tree contains a broader monophyletic group with the two species, *A. spinosus* and the rest of the Hybridus clade, the tree is not significantly longer (524 steps).

Finally, I wished to test the monophyly of the Galápagos clade, given that all of the Galápagos species and their close relatives occur in a single clade in the *G3PDH* tree, and the placement of various Galápagos clades within the genus is not highly supported in any other tree. When the nuclear tree or the chloroplast tree is constrained to keep all the Galápagos species and their close relatives in a single monophyletic group, the constrained trees are not significantly longer (1669 vs. 1666 steps, and 520 vs. 519 steps, respectively).

Trait Associations with Weediness

My calculations of Blomberg's K and the D statistic (using "phylo.d") in R did not support the hypothesis of phylogenetic signal in the traits agricultural weed status, problematic weed status, and agricultural weed rank, instead supporting a random distribution of weeds in the phylogeny for each metric of weediness. Therefore, I tested 12 morphological and ecological traits (five quantitative and seven qualitative) for associations with agricultural invasiveness using independent-samples t-tests and ANOVAs. Species' values recorded from the literature for each trait and for the agricultural invasiness metrics are given in Table 1.5. Of the five quantitative characters, four showed associations with weediness for at least one metric (Table 1.6). The number of GBIF cells occupied and the maximum elevation at which it grows were the quantitative traits most strongly associated with weediness, as weeds were found in more regions and grew at higher elevations for all three metrics. Maximum plant size was significant for two

of the weediness metrics, because problematic weeds were significantly larger than non-problematic and non-weedy plants. Average seed diameter was significant for a single metric, agricultural weed rank, but interestingly, this is because less problematic weeds (rank 2) had significantly larger seeds than either problematic weeds or non-weeds. *Amaranthus pumilus* has very large seeds relative to the rest of the genus (2.5 mm diameter), and this seed-size value was removed from the analyses as an outlier.

Some of the qualitative traits also showed strong associations with weediness (Table 1.6). Weeds are significantly more likely to grow in ruderal habitats and to have had their geographical range extended by humans, according to all three metrics. There are no weeds among the species that grow on beaches, which is significantly fewer than expected for two weediness metrics. Finally, weeds are more likely to occur in naturally disturbed habitats than non-weeds according to the agricultural weed status metric, but not using the other two metrics.

DISCUSSION

Phylogenetic Relationships and Topology Tests

Our phylogenetic analyses of relationships between species in the genus *Amaranthus* bolsters the suspected relationships between some species and clades based on morphology, and offers new insights into the relationships of other species, which were not obvious based on morphology. Three of the major clades roughly correspond to the three subgenera of Amaranthus recognized by Mosyakin and Robertson (1996) and Bayón (in review): the Eurasian/South African/Australian + South American clade corresponds to the subgenus Albersia, the Hybridus clade corresponds to the subgenus Amaranthus, and the Dioecious/Pumilus clade(s) correspond(s) to the subgenus Acnida (Figures 1.17 and 1.18). But

there are species in all three of these clades that were not predicted based on morphology, and no taxonomic authority has ever placed the Galápagos species and their relatives into a separate taxon or several separate taxa, instead lumping them into subgenus Albersia. (For possible relationships of species not sampled in this study, see Table 1.7.)

Biogeographical relationships among the species are also interesting, as the tree suggests that the genus probably originated and radiated first in the Americas, with only one clade giving rise to Old World species (Figures 1.19 and 1.20). I did not attempt to date my phylogeny, as the probable recent radiation of the group means that any molecular clock estimate of the genus' age would be dwarfed by standard error. However, Kadereit et al. (2003) used fossils to calibrate their estimates of the age of clades in the Amaranthaceae and Chenopodiaceae (both families are now placed into Amaranthaceae). They dated the root of the Chenopodiaceae at 65-56.5 million years old using two fossils, and with these plus another fossil at the crown of the Chenopodieae I clade, estimated a substitution rate of 2.8-4.1 synonymous substitutions per site per year for the chloroplast *rbcL* gene. Since there are 51 *rbcL* substitutions along the branches from the point of the Chenopodiaceae root to the genus *Amaranthus*, I can estimate the age of the *Amaranthus* root at 9.3 to 13.6 million years old. Even if this estimate is wildly inaccurate and the genus is as old as the Chenopodiaceae itself, the Old World species of *Amaranthus* almost definitely arose from long-distance dispersal, as South America, Africa, and Australia started to drift apart about 150 million years ago (Bortolotti and Principi, 2005). Furthermore, it appears from my phylogenies that a single long-distance dispersal event out of South America could have given rise to the entire ESA clade.

Relationships among species in the ESA+South American clade are generally very poorly resolved in my phylogenies, but the few well-resolved relationships have some precedent in the

taxonomic and phylogenetic literature on *Amaranthus*. Hunziker (1951) considered the South American species *A. persimilis, A. standleyanus, A. crispus*, and *A. cardenasianus* very similar morphologically, although he also thought *A. squamulatus* resembled *A. cardenasianus*, and believed that *A. kloosianus* was related to *A. urceolatus* rather than the former group. Bayón (in review) notes the close morphological similarity between *A. crispus* and *A. standleyanus*, and notes that *A. vulgatissimus* is similar to *A. deflexus*, although he places *A. cardenasianus* into subgenus Amaranthus rather than subgenus Albersia. Brenan (1981), contemplating introduced Amaranthus species in southern Africa, noticed that *A. deflexus, A. viridis*, and *A. muricatus* were similar, and Mosyakin and Robertson (2003) mention that *A. deflexus* and *A. muricatus* hybridize naturally.

No one seems to have predicted the apparent relationship between the Australian, Eurasian, and South African species, or any of the relationships between species within this group. Because of nomenclatural confusion surrounding *A. graecizans*, it is frequently mentioned in the taxonomic literature as similar to *A. albus* and *A. blitoides*, but Mosyakin and Robertston (2003) proposed that it was more closely related to Old World taxa with trimerous flowers, which is consistent with my results. Within the Australian species, Palmer (2009) says that *A. centralis* is most similar to *A. induratus*, which is echoed by Bayón (in review), but these species are not closely related in my trees, although the nuclear and chloroplast trees disagree on the placement of *A. centralis*. I did not include several South African species in my study (*A. schinzianus, A. dinteri*, and *A. capensis*), so the close relationship between *A. thunbergii* and *A. praetermissus* in my trees may be an artifact of sampling.

The Hybridus clade, on the other hand, has been the subject of many studies because of great interest in the origin of the grain amaranth species, *A. hypochondriacus* (from Mexico), *A.*

cruentus (from Guatemala), and *A. caudatus* (from the Andes) (Sauer, 1950). Sauer (1967) supported the hypothesis that *A. powellii*, *A. hybridus*, and *A. quitensis* were the respective progenitors of *A. hypochondriacus*, *A. cruentus*, and *A. caudatus*, while other authors have found support for Sauer's alternative hypothesis of a single or multiple origins of the grain amaranths from *A. hybridus* (Coons, 1977; Coons, 1978; Hauptli and Jain, 1984; Chan and Sun, 1997; Xu and Sun, 2001). Costea et al. (2001) completed a taxonomic treatment on the "*Amaranthus hybridus* species complex," including the grain amaranths, *A. hybridus*, *A.hybridus subsp. quitensis*, *A. powellii*, and *A. retroflexus*, supporting the recognition of the domesticated species as taxonomic entities separate from *A. hybridus*.

Several studies based on rapidly-evolving markers such as isozymes, RAPDs and microsatellites have produced polytomies of *A. hybridus* and the grain amaranths, with *A. quitensis* either inside or the sister taxon to this group if it is included in the study (Chan and Sun, 1997; Sun et al., 1999; Xu and Sun, 2001; Mallory, 2008); this pattern is essentially what my trees show. These same studies recovered the sister-lineage relationship of *A. powellii* and/or *A. retroflexus* (shown in my nuclear tree) to this core *A. hybridus* group. Mosyakin and Robertson (2003) note that *A. wrightii* is closely related to *A. retroflexus*, which is borne out in my study, and also hypothesizes that *A. spinosus* is probably the sister taxon to subgenus Amaranthus, which my nuclear tree supports. Chan and Sun (1997) included *A. acutilobus* in their isozyme and RAPD phylogenetic study and inferred that it fell within the *A. hybridus* clade, which is also consistent with my results. Future phylogenetic work within the Hybridus clade should take into account the probable rampant hybridization between the domesticated species, *A. hybridus*, and *A. quitensis* (since the latter two species have undoubtedly been associated weeds of *Amaranthus* crop fields since domestication; Sauer, 1950; 1967), and should use

phylogenetic estimation methods that account for reticulate evolution.

Sauer (1967) hypothesized from the cytological work of Grant (1959) that *A. spinosus* probably hybridized with a species of the *A. hybridus* complex to create the allotetraploid *A. dubius*. My chloroplast tree strongly supports *A. dubius* as the sister lineage to *A. spinosus*, which leads me to believe that *A. spinosus* or the lineage that led to *A. spinosus* was its maternal parent. In the nuclear tree, *A. dubius* is supported as belonging to the core Hybridus group, although I cannot tell which of the species in the complex was its paternal parent: single nuclear genes support one allele of each accession with *A. spinosus* and the other allele in the core Hybridus group. The placement of *A. dubius* in the Hybridus clade was anticipated genetically by Chan and Sun (1997).

Finally, the placement of *A. palmeri* and *A. watsonii* has been unclear in the previous literature. Their close relationship to each other is clear based on morphology (Standley, 1914; Brenan, 1961). Mosyakin and Robertson's (1996) taxonomic treatment included all the dioecious *Amaranthus* species in subgenus Acnida, even though the author recognized that the group was "artificial and polyphyletic" (Mosyakin and Robertson, 2003). There were several previous indications that *A. palmeri* may be related to the Hybridus clade: Franssen et al., (2001) noticed that the pollen morphology of *A. palmeri* was unlike that of the other dioecious *Amaranthus* species sampled and more closely resembled that of the monoecious species. Chan and Sun (1997) placed *A. palmeri* as the sister lineage to their *A. hybridus* clade with isozyme and RAPD data, Wassom and Tranel (2005) placed *A. palmeri* and *A. spinosus* together based on AFLP data, and Riggins et al. (2010) placed *A. palmeri* and *A. spinosus* together and as the sister group to the Hybridus clade based on the *ALS* gene.

The present study found strongly-supported disagreement between the nuclear

chloroplast datasets in the placement of *A. palmeri* + *A. watsonii* (further upheld by Templeton tests), suggesting a possible ancient chloroplast capture event from the lineage leading to the other dioecious species (see Rieseberg and Soltis, 1991; Rieseberg et al., 1996; and Tsitrone et al., 2003 for reviews of chloroplast capture in plant phylogenies and conditions that promote capture). This appears more likely than a hybridization event in which nuclear material from both hybridizing species was retained, as none of the four nuclear genes support the chloroplast tree's placement of *A. palmeri* + *A. watsonii* in the Dioecious/Pumilus clade. It is curious that *A. pumilus* is the most closely related species to *A. palmeri* + *A. watsonii* according to the chloroplast tree, because they are native to opposite ends of a continent: *A. pumilus* is an endangered beach specialist endemic to the Atlantic coast of the U.S., and *A. palmeri* and *A. watsonii* are both from the southwestern U.S. The fact that *A. palmeri* and *A. watsonii* are the only dioecious species placed outside of the Dioecious/Pumilus clade by the nuclear tree also suggests that dioecy in *Amaranthus* might be a trait encoded or influenced strongly by the chloroplast. However, dioecy is dominant over the monoecious condition in crosses of monoecious species with *A. tuberculatus*, regardless of the direction of the cross, which implies a nuclear element in breeding-system determination (Murray, 1940; Trucco et al., 2006).

The apparent inclusion of *A. pumilus* in the Dioecious/Pumilus clade was anticipated by one previous study. Nolan et al. (2010) studied the population genetics and phylogenetic relationships of *A. pumilus* using ISSRs, and found that *A. arenicola* was weakly grouped with *A. pumilus* by neighbor-joining and Bayesian inference, although these analyses did not group the other sampled dioecious species with this clade. No other authors have put forward hypotheses about the relationship of *A. pumilus*, because of its morphological distinctiveness in the genus. This federally endangered monoecious species has larger seeds than do any other *Amaranthus*

species, and the entire plant is fleshy (Mosyakin and Robertson, 2003).

Incomplete lineage sorting leads to problems with recovering the species tree from single-gene trees (Degnan and Rosenberg, 2009). In my analyses, the topology of the Dioecious/Pumilus clade is different in my phylogenies based on different nuclear genes. Concatenation of genes may lead to an incorrect species-tree phylogeny when gene trees differ and molecular models of evolution are different for each gene (Degnan and Rosenberg, 2009), but my concatenated nuclear dataset is partitioned to account for this. Increased within-species sampling can improve the likelihood of estimating the true species tree for shallower phylogenies (Maddison and Knowles, 2006). A number of new methods for estimating species' trees in the presence of incomplete lineage sorting are becoming available (Degnan and Rosenberg, 2009); an in-depth study of relationships in the Dioecious/Pumilus clade would ideally include more within-species sampling and would test some of the new methods for congruence. It should be noted that the strongly supported non-monophyly of *A. tuberculatus* in the concatenated nuclear tree (as opposed to monophyly in the chloroplast tree) may not be an artifact of incomplete lineage sorting, but a correct reflection of evolutionary history. *Amaranthus tuberculatus* was previously considered two largely allopatric species based on morphology (Sauer, 1967; Pratt and Clark, 2001), and the sample of *A. tuberculatus* from west of the Mississippi River is placed phylogenetically with other western dioecious species, whereas the sample from east of the Mississippi River is grouped with eastern North American dioecious species. The two "species" are now considered varieties by some authors (Costea et al. 2005), and it is possible that they or their ancestral taxa might have originated separately from different dioecious groups and subsequently coalesced into one species through hybridization.

The Galápagos clades in my trees support the relationships of the three to four endemic

or native Galápagos *Amaranthus* species with North and South American species. Some of these relationships were predicted based on morphology: Eliasson (1985; 1987) notes that the Galápagos species *A. anderssonii* and the Caribbean *A. berlandieri* (=*A. polygonoides*) are virtually indistinguishable morphologically. However, Eliasson also thought that *A. anderssonii* and *A. squamulatus*, another Galápagos native that also occurs in mainland Ecuador, were closely related, which is not supported in my trees. Instead, *A. squamulatus* forms a separate clade with a pair of species from North America (*A. fimbriatus*) and South America (*A. urceolatus*), and oddly enough, it appears to be the sister lineage to this group, rather than being more closely related to the South American species. Its relationship to *A. urceolatus* was predicted by Eliasson (1987), but no authority has previously linked the desert species *A. fimbriatus* to this group. Furthermore, the hybrid North American species *A. x tucsonensis* is placed with this clade in the nuclear tree, which is unexpected, as its authority Henrickson (1999) eliminated *A. fimbriatus* as a parent based on morphology.

Several authorities have recognized the similarity of *A. albus, A. blitoides*, and *A. californicus* (Mosyakin and Robertson, 2003; Bayón, in review). *Amaranthus albus* and *A. blitoides* were placed together in a neighbor-joining tree based on *ALS* gene sequence data by Riggins et al. (2010). The only author to connect the Galápagos species *A. sclerantoides* to this group was Hunziker (1965), who placed nine species in a group based on their axillary inflorescences, 1-5 tepals, and 1-5 stamens, and included all four of the species in this clade. However, he also included *A. looseri* and *A. acutilobus*, which are supported in my study as belonging to the paraphyletic South American group subtending the ESA clade, and the Hybridus clade, respectively. Finally, the putative new species, which is highly supported as sister to the Caribbean *A. crassipes*, needs morphological description to determine if it is truly

distinct from this species. It is known from only two populations on two islands in Galápagos, and could thus represent a persistent, early introduction rather than a speciation event.

Templeton tests are unable to rule out the possibility that all of the Galápagos species and their close relatives, which appear as three strongly-supported clades in the nuclear tree and four clades plus an extraneous species in the chloroplast tree, are actually a single monophyletic group. Three to four Galápagos colonization events from a single group of *Amaranthus* and none from any of the other three clades in the genus would imply that successful colonization of the Galapagos islands involves a non-random set of ancestral traits. More phylogenetic work to resolve the relationships among the major clades of the genus might resolve this point. Furthermore, the connection of the Galápagos species to Caribbean and southwestern North American relatives is congruent with the discovery that a number of endemic Galápagos species previously assumed to be closely tied to nearby South America (Porter, 1979) in fact originated in the Caribbean, Central America, "Tropical America," or southwestern North America (Tye and Francisco-Ortega, 2011). The endemic Galápagos *Amaranthus* species are probably dispersed internal or externally (in mud) by birds, and many Galápagos birds are migratory and travel thousands of miles each year (Porter, 1983). Rare bird dispersal of *Amaranthus* to the archipelago could explain the Galápagos biogeography seen in this genus.

Trait Associations with Weediness

The lack of phylogenetic signal in any of the metrics of agricultural invasiveness supports the idea of a lack of phylogenetic constraint in the evolution of weeds in *Amaranthus*, and/or homoplasy in weedy traits (see Figures 1.21 and 1.22). The fact that I do find some traits associated with weediness in the genus suggests that there is at least some parallel evolution of

the same traits in the genus, but given the phylogenetic distinctiveness of some of the weeds, the weed species are probably also successful in crop fields for different reasons, at least to some degree. For instance, breeding system (dioecy vs. monoecy) and utricle dehiscence (dehiscent vs. indehiscent) are not significantly associated with weediness in *Amaranthus*, and yet these characteristics have substantial bearing on genetic diversity and seed dispersal, respectively.

For my analyses, I collected trait data from the literature (as in Jenkins and Keller, 2011 for invasive and non-invasive *Silene* species). This is an ad-hoc analysis and is not meant to conclusively pinpoint the morphological and ecological traits associated with agricultural invasiveness in *Amaranthus*. Ideally, common-garden experiments comparing the traits of weedy and non-weedy species within agricultural and non-agricultural ecosystems would be conducted (i.e. Hodgins and Rieseberg, 2011, except at the interspecific level), similar to what has previously been done for invasive plants of natural ecosystems (Burns, 2004; Schlaepfer et al., 2010; Fenesi et al., 2011). This would allow more rigorous testing of some of the hypotheses generated by my phylogenetic analysis in *Amaranthus* and other groups, as well as the opportunity to examine traits associated with competition that are seldom recorded in species' descriptions (i.e., relative growth rate, germination rate and timing, flowering time and duration, fecundity, and plasticity in all these characters; as in Muth and Pigliucci, 2006; Schlaepfer et al., 2010; Van Kleunen et al., 2011; for invasive plants). Such experiments were not done as part of this study due to the difficulty of obtaining seeds for many *Amaranthus* species, and the restrictions on the use of seeds by the country of origin for other species.

Qualifications aside, my analyses did uncover some intriguing associations between weediness and other traits in the genus. The association of agricultural invasiveness with the number of GBIF cells occupied was not unexpected, given that geographical range size is often

associated with invasiveness or weediness, although this is not necessarily the case (Williamson and Fitter, 1996; Lososová et al., 2008). An interesting follow-up to this finding would be an examination of the percent of land under intensive agricultural production within the geographical range of the species: it is possible that some species have simply had more opportunity to invade agricultural fields because of where they are native. An important qualification on the use of GBIF records to estimate species' ranges is that the database includes all records, whether native or introduced to the region (and some records are probably misidentified). The amount of area originally occupied by the species is very hard to estimate, particularly for species that have been extensively moved around by human activity. I have tried to account for these range extensions to some degree with the binary character "geographical range expanded by humans," with data derived from the literature; unsurprisingly, this character is associated with agricultural weediness.

The association of maximum plant size with agricultural invasiveness is also unsurprising, as shading is an important aspect of competition among plants (Eriksen et al., 2012). The nonsignificant results for maximum leaf length were less expected, as were the results for average seed size: it is unclear why non-problematic weeds would have larger seeds than either non-weeds or problematic weeds. Seed size can be negatively or positively associated with invasiveness (Hamilton et al., 2005; Lloret et al., 2005), and there could conceivably be an advantage to having smaller seeds in agricultural ecosystems, especially since smaller seeds have been shown to have longer dormancy in the soil (Venable and Brown, 1988; Thompson et al., 1993; but see Leishman et al., 2000).

Many of the agricultural weed species in *Amaranthus* are also found in ruderal (waste ground) habitats, and this is borne out as statistically significant in my analyses. It is interesting

that natural disturbance was not strongly associated with weediness in the group, and also very interesting that *Amaranthus* weeds are never found on beaches, which are a quintessential naturally disturbed environment. These findings may be specific to the genus, but have not (to my knowledge) been examined in other groups. Observations of the endemic Galápagos *Amaranthus* species show that populations of the littoral species *A. sclerantoides* and *A. anderssonii* have disappeared from occupied areas of the islands, possibly due to human population pressure (K. Waselkov, *pers. obs.*). This suggests that adaptation to natural or littoral disturbance versus human disturbance may require different morphological traits or physiological tolerances in the group. The finding that weeds can grow at higher maximum elevations than non-weeds, and problematic weeds can grow at higher elevations than non-problematic weeds, is also unexpected and may be specific to *Amaranthus*, but should be explored in other groups.

In conclusion, I have presented here some strongly-supported relationships of clades and species in the genus *Amaranthus*, and some initial associations of traits in the genus with agricultural invasiveness. This study could be the starting point for investigations into relationships between and within subgenera of *Amaranthus*, further testing of biogeographic hypotheses within the genus, and the study of the evolution and underlying genetics of breeding systems in the group. In addition, I hope to prompt more investigation into the understudied realm of general traits of agricultural weeds, given some of the surprising findings in my preliminary study.

LITERATURE CITED

BAKER, H. G. 1974. The evolution of weeds. *Annual Review of Ecology and Systematics* 5: 1-24.

BAYÓN, N. D. In review. Revisión taxonómica de las especies monoicas de *Amaranthus* L. (Amaranthaceae): Amaranthus subg. Albersia y Amaranthus subg. Amaranthus. *Annals of the Missouri Botanical Garden*. Used with permission.

BENECH-ARNOLD, R. L., R. A. SÁNCHEZ, F. FORCELLA, B. C. KRUK, AND C. M. GHERSA. 2000. Environmental control of dormancy in weed seed banks in soil. *Field Crops Research* 67: 105-122.

BORTOLOTTI, V., AND G. PRINCIPI. 2005. Tethyan ophiolites and Pangea break-up. *The Island Arc* 14: 442-470.

BRÄNDLE, M., J. STADLER, S. KLOTZ, AND R. BRANDL. 2003. Distributional range size of weedy plant species is correlated to germination patterns. *Ecology* 84: 136–144.

BRENAN, J. P. M. 1961. *Amaranthus* in Britain. *Watsonia* 4: 261--280.

BRENAN, J. P. M. 1981. The genus *Amaranthus* in Southern Africa. *Journal of South African Botany* 47: 451-492.

BURNS, J.H. 2004. A comparison of invasive and non-invasive dayflowers (Commelinaceae) across experimental nutrient and water gradients. *Diversity and Distributions* 10: 387–397.

CHAN, K. F. AND M. SUN. 1997. Genetic diversity and relationships detected by isozime and RAPD analysis of crop and wild species of *Amaranthus*. Theoretical and Applied Genetics 95: 865-873.

CHANEY, L., AND R.S. BAUCOM. 2012. The evolutionary potential of Baker's weediness traits in the common morning glory, *Ipomoea purpurea* (Convolvulaceae). *American Journal of Botany* 99: 1524–1530.

COONS, M. P. 1977. The status of *Amaranthus hybridus* L. in South America. *Ciencia y Naturaleza* 18: 80-87.

COONS, M. P. 1978. The status of *Amaranthus hybridus* L. in South America: The taxonomic problem. *Ciencia y Naturaleza* 19: 66-71.

COSTEA, M., A. SANDERS AND G. WAINES. 2001a. Preliminary results toward a revision of the *Amaranthus hybridus* complex (Amaranthaceae). *Sida* 19(4): 931-974.

COSTEA, M., S. E. WEAVER, AND F. J. TARDIF. 2005. The biology of invasive alien plants in

Canada. 3. *Amaranthus tuberculatus* (Moq.) Sauer var. *rudis* (Sauer) Costea & Tardif. *Canadian Journal of Plant Science* 85: 507–522.

DAEHLER, C. C. 1998. The taxonomic distribution of invasive angiosperm plants: ecological insights and comparison to agricultural weeds. *Biological Conservation* 84: 167–180.

DEGNAN, J.H., AND N. A ROSENBERG. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution* 24: 332–340.

DOYLE, J. J. AND J. L. DOYLE. 1990. Isolation of plant DNA from fresh tissue. *Focus* 12: 13–15.

ELIASSON, U. H. 1985. Identity and taxonomic affinity of some members of the Amaranthaceae from the Galápagos Islands. *Botanical Journal of the Linnaean Society* 91: 415-433.

ELIASSON, U. H. 1987. Amaranthaceae. *In* G. Harling and L. Andersson [eds.], *Flora of Ecuador* 44: 1-138.

ELIASSON, U. H. 1988. Floral morphology and taxonomic relations among the genera of Amaranthaceae in the New World and the Hawaiian Islands. *Botanical Journal of the Linnaean Society* 96: 235-283.

ERIKSEN, R. L., T. DESRONVIL, J. L. HIERRO, AND R. KESSELI. 2012. Morphological differentiation in a common garden experiment among native and non-native specimens of the invasive weed yellow starthistle (*Centaurea solstitialis*). *Biological Invasions* 14: 1459-1467.

FELSENSTEIN, J. 1985. Phylogenies and the comparative method. *The American Naturalist* 125: 1–15.

FELSENSTEIN J. 2005. PHYLIP (phylogeny inference package). Distributed by the author. University of Washington Department of Genome Sciences, Seattle, Washington, USA.

FENESI, A., T. RÉDEI, AND Z. BOTTA-DUKÁT. 2011. Hard traits of three *Bromus* species in their source area explain their current invasive success. *Acta Oecologica* 37: 441–448.

FRANSSEN, A.S., D.Z. SKINNER, K. AL-KHATIB, AND M.J. HORAK. 2001. Pollen morphological differences in *Amaranthus* species and interspecific hybrids. *Weed Science* 49: 732–737.

GRANT, W. F. 1959. Cytogenetic studies in *Amaranthus*. II. Natural interspecific hybridization between *Amaranthus dubius* and *A. spinosus*. *Canadian Journal of Botany* 37: 1063--1070.

GRENIER, J. C. M. AND D. A. GODRON. 1848. Flore de France, V. 3, pp. 3-6. Libraire de académie impérial de médecine, J.-B. Bailliére, Paris, France.

HALL, T. 2005. BioLign 4.0.2.6 alignment and multiple contig editor. Ibis Biosciences, Carlsbad,

California, USA.

HAMILTON, M. A., B.R. MURRAY, M.W. CADOTTE, G.C. HOSE, A.C. BAKER, C.J. HARRIS, AND D. LICARI. 2005. Life-history correlates of plant invasiveness at regional and continental scales. *Ecology Letters* 8: 1066–1074.

HAUPTLI, H., AND S. JAIN. 1984. Allozyme variation and evolutionary relationships of grain amaranths (*Amaranthus* spp.). *Theoretical and Applied Genetics* 69: 153-165.

HENRICKSON, J. 1999. Studies in New World *Amaranthus* (Amaranthaceae). *Sida* 18: 783--807.

HENRICKSON, J. 2004. *Amaranthus acanthobracteatus* (Amaranthaceae). *Sida* 21: 11-17.

HILLIS, D. M., AND J. J. BULL. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology* 42: 182-192.

HODGINS, K. A., AND L. RIESEBERG. 2011. Genetic differentiation in life-history traits of introduced and native common ragweed (*Ambrosia artemisiifolia*) populations. *Journal of Evolutionary Biology* 24: 2731-2749.

HUNZIKER, A. T. 1951. Cuatro nuevas especies Sudamericanas de *Amaranthus*. *Boletín de la Sociedad Argentina de Botaníca* 4: 133-143.

HUNZIKER, A. T. 1965. Estudios sobre *Amaranthus*. V. Revisión de las especies americanas con 1 a 4 estambres, 1 a 5 tépalos, e inflorescencias exclusivamente axilares. *Kurtziana* 2: 27-52.

HUNZIKER, A. T. 1966. Estudios sobre *Amaranthus*. VI. Notas taxonómicas sobre especies sudamericanas. *Kurtziana* 3: 201-214.

JENKINS, C., AND S. R. KELLER. 2011. A phylogenetic comparative study of preadaptation for invasiveness in the genus *Silene* (Caryophyllaceae). *Biological Invasions* 13: 1471-1486.

KADEREIT, G., T. BORSCH, K. WEISING, AND H. FREITAG. 2003. Phylogeny of Amaranthaceae and Chenopodiaceae and the evolution of C_4 photosynthesis. *International Journal of Plant Sciences* 164: 959–986.

KEMBEL, S. W., P. D. COWAN, M. R. HELMUS, W. K. CORNWELL, H. MORLON, D. D. ACKERLY, S. P. BLOMBERG, AND C. O. WEBB. 2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26:1463-1464.

KUNTH, K. S. 1838. Flora Berolinensis, V. 2, pp. 144-145. Duncker and Humblot, Berlin, Germany.

LANOUE, K. Z., P. G. WOLF, S. BROWNING & E. E. HOOD. 1996. Phylogenetic analysis of restriction-site variation in wild and cultivated *Amaranthus* species (Amaranthaceae).

Theoretical and Applied Genetics 93: 722-732.

LEISHMAN, M. R., I. J. WRIGHT, A. T. MOLES, AND M. WESTOBY. 2000. Chapter 2: The evolutionary ecology of seed size. *In* M. Fenner [ed.], Seeds: The Ecology of Regeneration in Plant Communities, 2nd edition, pp. 31-57. CABI Publishing, New York, USA.

LINNAEUS, C. 1753. Species Plantarum, V. 2, pp. 989. Laurentius Salvius, Stockholm, Sweden.

LLORET, F., F. MÉDAIL, G. BRUNDU, I. CAMARDA, E. MORAGUES, J. RITA, P. LAMBDON, AND P. E. HULME. 2005. Species attributes and invasion success by alien plants on Mediterranean islands. *Journal of Ecology* 93: 512-520.

LOSOSOVÁ, Z., M. CHYTRÝ, AND I. KÜHN. 2008. Plant attributes determining the regional abundance of weeds on central European arable land. *Journal of Biogeography* 35: 177-187.

MADDISON, W.P., AND L. L. KNOWLES. 2006. Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology* 55: 21-30.

MALLORY, M.A., R. V. HALL, A.R. MCNABB, D.B. PRATT, E.N. JELLEN, AND P.J. MAUGHAN. 2008. Development and characterization of microsatellite markers for the grain amaranths. *Crop Science* 48: 1098–1106.

MOSYAKIN, S. L. & K. R. ROBERTSON. 1996. New infrageneric taxa and combinations in *Amaranthus* L. (Amaranthaceae). *Annales Botanici Fennici* 33: 275-281.

MOSYAKIN, S., AND K. R. ROBERTSON. 2003. *Amaranthus*. *In* Flora of North America Editorial Committee [eds.], Flora of North America North of Mexico, Vol. 4. Magnoliophyta: Caryophyllidae, Pt. 1: 410-435. Oxford University Press, New York, USA.

MÜLLER, K., AND T. BORSCH. 2005. Phylogenetics of Amaranthaceae based on *matK/trnK* sequence data: Evidence from parsimony, likelihood, and Bayesian analyses. *Annals of the Missouri Botanical Garden* 92: 66–102.

MURRAY, J.M. 1940. The genetics of sex determination in the family Amaranthaceae. *Genetics* 25: 409–431.

MUTH, N. Z. AND M. PIGLIUCCI. 2006. Traits of invasives reconsidered: phenotypic comparisons of introduced invasive and introduced noninvasive plant species within two closely related clades. *American Journal of Botany* 93: 188-196.

NOLAN, C., A. NOYES, A. BENNETT, R. HUNTER, AND K.L. HUNTER. 2010. Inter simple sequence repeats (ISSR) reveal genetic variation among Mid-Atlantic populations of threatened *Amaranthus pumilus* and phylogenetic relationships. *Castanea* 75: 506–516.

NYLANDER, J. A. A. 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

ORME, D., R. FRECKLETON, G. THOMAS, T. PETZOLDT, S. FRITZ, N. ISAAC, AND W. PEARSE. 2011. Caper: Comparative Analyses of Phylogenetics and Evolution in R. R package version 0.5.

PALMER, J. 2009. A conspectus of the genus *Amaranthus* L. (Amaranthaceae) in Australia. Nuytsia 19: 107-128.

PARADIS E., J. CLAUDE, AND K. STRIMMER. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20: 289-290.

PARK, Y-J., K. NEMOTO, T. NISHIKAWA, K. MATSUSHIMA, M. MINAMI, AND M. KAWASE. 2010. Waxy strains of three amaranth grains raised by different mutations in the coding region. *Molecular Breeding* 25: 623-635.

PERRINS, J., M. WILLIAMSON, AND A. FITTER. 1992. A survey of differing views of weed classification: implications for regulation of introductions. *Biological Conservation* 60: 47-56.

PORTER , D. M. (1979). Endemism and evolution in Galápagos Islands vascular plants. *In* D. Bramwell [ed.], Plants and Islands, pp. 225 – 256. Academic Press, London, UK.

PORTER, D. M. 1983. Vascular plants of the Galápagos: Origins and dispersal. *In* R. I. Bowman, M. Berson, and A. E. Leviton [eds.], Patterns of evolution in Galápagos organisms, pp. 33-96. Pacific Division, AAAS, San Francisco, USA.

PRATT, D. B., AND L. G. CLARK. 2001. *Amaranthus rudis* and *A. tuberculatus*, one species or two? *Journal of the Torrey Botanical Society* 128: 282-296.

PYŠEK, P., AND D.M. RICHARDSON. 2007. Traits associated with invasiveness in alien plants: Where do we stand? *In* W. Nentwig [ed.], Biological Invasions, Ecological Studies 193, pp. 97–126. Springer-Verlag, Berlin & Heidelberg, Germany.

RAMBAUT, A. 2012. FigTree v1.4.0, http://tree.bio.ed.ac.uk/software/figtree/. Accessed May 25, 2013.

RIESEBERG, L. H., AND D. E. SOLTIS. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. *Evolutionary Trends in Plants* 5: 65-84.

RIESEBERG, L. H., J. WHITTON, AND C. R. LINDER. 1996. Molecular marker incongruence in plant hybrid zones and phylogenetic trees. *Acta Botanica Neerlandica* 45: 243-262.

RIGGINS, C. W., Y. PENG, C. N. STEWART JR., AND P. J. TRANEL. 2010. Characterization of de novo transcriptome for waterhemp (*Amaranthus tuberculatus*) using GS-FLX 454

pyrosequencing and its application for studies of herbicide target-site genes. *Pest Management Science* 66: 1042-1052.

ROBERTSON, K. 1981. The genera of Amaranthaceae in the Southeastern United States. *Journal of the Arnold Arboretum* 62: 267-314.

RONQUIST, F. AND J.P. HUELSENBECK. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.

SAGE, R.F., T.L. SAGE, R.W. PEARCY, AND T. BORSCH. 2007. The taxonomic distribution of C_4 photosynthesis in Amaranthaceae sensu stricto. *American Journal of Botany* 94: 1992–2003.

SAUER, J. D. 1950. The grain amaranths: a survey of their history and classification. *Annals of the Missouri Botanical Garden* 37: 561-632.

SAUER, J. D. 1955. Revision of the dioecious amaranths. *Madroño* 13: 5-46.

SAUER, J.D. 1967. The grain amaranths and their relatives: a revised taxonomic and geographic survey. *Annals of the Missouri Botanical Garden* 54: 103–137.

SCHLAEPFER, D. R., M. GLÄTTLI, M. FISCHER, AND M. VAN KLEUNEN. 2010. A multi-species experiment in their native range indicates pre-adaptation of invasive alien plant species. *New Phytologist* 185: 1087-1099.

SHAW, J., E. B. LICKEY, J. T. BECK, S. B. FARMER, W. LIU, J. MILLER, K. C. SIRIPUN, C. T. WINDER, E. E. SCHILLING, AND R. L. SMALL. 2005. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany* 92: 142-166.

SIKES, D. S. AND P. O. LEWIS. 2001. beta software, version 1. PAUPRat: PAUP* implementation of the parsimony ratchet. Distributed by the authors. Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, USA.

STAMATAKIS, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688-2690.

STANDLEY, P.C. 1914. New or notable species of *Amaranthus*. *Bulletin of the Torrey Botanical Club* 41: 505–510.

STRAND, A. E., J. LEEBENS-MACK, AND B. G. MILLIGAN. 1997. Nuclear DNA-based markers for plant evolutionary biology. *Molecular Ecology* 6: 113-118.

SUN, M., H. CHEN, AND F.C. LEUNG. 1999. Low-Cot DNA sequences for fingerprinting analysis of germplasm diversity and relationships in *Amaranthus*. *Theoretical and Applied Genetics* 99: 464–472.

SWOFFORD, D. L. 1998. PAUP*: phylogenetic analysis using parsimony (*and other methods). Version 4.0. Sinauer, Sunderland, MA.

SUTHERLAND, S. 2004. What makes a weed a weed: life history traits of native and exotic plants in the USA. *Oecologia* 141: 24–39.

TAMURA, K., D. PETERSON, N. PETERSON, G. STECHER, M. NEI, AND S. KUMAR. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* 28: 2731-2739.

TEMPLETON, A. R. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the humans and apes. *Evolution* 37: 221-244.

TEMPLETON, A. R. 2006. Population Genetics and Evolutionary Theory. John Wiley and Sons Inc., Hoboken, New Jersey.

THOMPSON, K., S.R. BAND, AND J.G. HODGSON. 1993. Seed size and shape predict persistence in soil. *Functional Ecology* 7: 236–241.

TRUCCO, F., T. TATUM, K.R. ROBERTSON, A.L. RAYBURN, AND P.J. TRANEL. 2006. Characterization of waterhemp (*Amaranthus tuberculatus*) x smooth pigweed (*A. hybridus*) F_1 hybrids. *Weed Technology* 20: 14–22.

TSITRONE, A., M. KIRKPATRICK, AND D. A. LEVIN. 2003. A model for chloroplast capture. *Evolution* 57: 1776-1782.

TYE, A., AND J. FRANCISCO-ORTEGA. 2011. Origins and evolution of Galápagos endemic vascular plants. *In* D. Bramwell and J. Caujapé-Castells, [eds.], The Biology of Island Floras, 89-153. Cambridge University Press, Cambridge, UK.

VAN KLEUNEN, M., AND S. D. JOHNSON. 2007. South African Iridaceae with rapid and profuse seedling emergence are more likely to become naturalized in other regions. *Journal of Ecology* 95: 674-681.

VAN KLEUNEN, M., J. C. MANNING, V. PASQUALETTO, AND S. D. JOHNSON. 2008. Phylogenetically independent associations between autonomous self-fertilization and plant invasiveness. *The American Naturalist* 171: 195-201.

VAN KLEUNEN, M., D. R. SCHLAEPFER, M. GLAETTLI, AND M. FISCHER. 2011. Preadapted for invasiveness: do species traits or their plastic response to shading differ between invasive and non-invasive plant species in their native range? *Journal of Biogeography* 38: 1294-1304.

VENABLE, D. L., AND J. S. BROWN. 1988. The selective interactions of dispersal, dormancy, and seed size as adaptations for reducing risk in variable environments. *The American Naturalist* 131: 360-384.

WASSOM, J.J., AND P.J. TRANEL. 2005. Amplified fragment length polymorphism-based genetic relationships among weedy *Amaranthus* species. *The Journal of Heredity* 96: 410–416.

WILLIAMSON, M. H., AND A. FITTER. 1996. The characters of successful invaders. *Biological Conservation* 78: 163–170.

XU, F., AND M. SUN. 2001. Comparative analysis of phylogenetic relationships of grain amaranths and their wild relatives (*Amaranthus*; Amaranthaceae) using internal transcribed spacer, amplified fragment length polymorphism, and double-primer fluorescent intersimple sequence repeat. *Molecular Phylogenetics and Evolution* 21: 372–87.

TABLES AND FIGURES

Table 1.1. Species of *Amaranthus* included in the phylogenetic reconstruction. "Authority" is the taxonomic authority for the name. "Subgenus Mosyakin and Robertson" lists the subgenus the species falls into according to the 1996 classification of Mosyakin and Robertson: a question mark after the classification means that the authors did not specifically mention the species in their article, and I used their morphological criteria to place it into a subgenus and section. "Subgenus Bayón" lists the subgenus the species falls into according to the revision of the monoecious species by Bayón (in review). Geographical origin is the native range of the species; in the case of weeds, this is the consensus native range in the literature.

Genus	**Species**	**Subspecies**	**Authority**	**Subgenus Mosyakin and Robertson**	**Subgenus Bayón**	**Geographical origins: Continent(s)**	**Geographical origins: Areas within continent(s)**
Amaranthus	*acanthochiton*		J.D. Sauer	Acnida sect. Acanthochiton	n/a	North America	Southwest US, Chihuahua
Amaranthus	*acutilobus*		Uline & Bray	Albersia sect. Pentamorion	Albersia	Mexico	Southern Mexico
Amaranthus	*albus*		L.	Albersia sect. Pyxidium	Albersia	North America	Western U.S.
Amaranthus	*anderssonii*		Howell	Albersia sect. Pentamorion?	Albersia	South America	Galápagos Islands
Amaranthus	*arenicola*		I.M. Johnson	Acnida sect. Saueranthus	n/a	North America	US Great Plains
Amaranthus	*australis*		(A. Gray) J.D. Sauer	Acnida sect. Acnida	n/a	North America/Central America/South America	Southeastern US, eastern Mexico, West Indies, northern South America
Amaranthus	*blitoides*		S. Watson	Albersia sect. Pyxidium	Albersia	North America	Central and part of Eastern US
Amaranthus	*blitum*	subsp. blitum	L.	Albersia sect. Blitopsis	Albersia	Eurasia	Eurasia
Amaranthus	*blitum*	subsp. emarginatus	(Moq. ex Uline & Bray) Carretero	Albersia sect. Blitopsis (as *A. emarginatum*)	Albersia (as subsp. polygonoides)	Eurasia	
Amaranthus	*blitum*	subsp. oleraceus	(L.) Costea	n/a	Albersia	Eurasia	
Amaranthus	*blitum*	subsp. emarginatus var. pseudogracilis	(Thell.) Costea	n/a	Albersia (as subsp. polygonoides var. pseudogracilis)	Eurasia	
Amaranthus	*californicus*		(Moq.) S. Watson	Albersia sect. Pyxidium?	Albersia	North America	Western U.S. and Canada
Amaranthus	*cannabinus*		(L.) J.D. Sauer	Acnida sect. Acnida	n/a	North America	US Atlantic coast
Amaranthus	*cardenasianus*		Hunz.	Albersia sect. Pyxidium?	Amaranthus	South America	Argentina, Bolivia, and Peru

Genus	Species	Subspecies	Authority	Subgenus Mosyakin and Robertson	Subgenus Bayón	Geographical origins: Continent(s)	Geographical origins: Areas within continent(s)
Amaranthus	*caudatus*		L.	Amaranthus sect. Amaranthus	Amaranthus	South America	Andean highlands
Amaranthus	*centralis*		J. Palmer & Mowatt	Albersia sect. Pentamorion?	Albersia	Australia	Central and northwestern Australia
Amaranthus	*clementii*		Domin	Albersia sect. Pentamorion	Albersia	Australia	Western Australia
Amaranthus	*cochleitepalus*		Domin	Albersia sect. Pentamorion?	Albersia	Australia	Western, northern, and Queensland regions of Australia
Amaranthus	*crassipes*		Schltdl.	Albersia sect. Pentamorion	Albersia	North America/Central America/South America	Gulf of Mexico and surrounding areas, to northern South America
Amaranthus	*crispus*		(Lespinasse & Thévenau) A. Braun ex J. M. Coulter & S. Watson	Albersia sect. Pentamorion	Albersia	South America	Argentina, Chile, Uruguay
Amaranthus	*cruentus*		L.	Amaranthus sect. Amaranthus	Amaranthus	North America/Central America	Guatemala and Mexico
Amaranthus	*cuspidifolius*		Domin	Albersia sect. Pentamorion?	Albersia	Australia	Central and western Australia
Amaranthus	*deflexus*		L.	Albersia sect. Blitopsis	Albersia	South America	South American pampas
Amaranthus	*dubius*		Mart. ex Thell.	Amaranthus sect. Dubia	Amaranthus	Central and South America	West Indies and northern South America
Amaranthus	*fimbriatus*		(Torr.) Benth. ex S. Watson	Albersia sect. Pyxidium?	Amaranthus	North America	Southwest US, northern Mexico
Amaranthus	*floridanus*		(S. Watson) J.D. Sauer	Acnida sect. Acnida	n/a	North America	Florida
Amaranthus	*graecizans*		L.	Albersia sect. Pyxidium	Albersia	Eurasia	Mediterranean, Asia, north Africa
Amaranthus	*graecizans*	subsp. aschersonianus	(Thell.) Costea, Brenner, & Tardif	n/a	n/a	Eurasia	
Amaranthus	*graecizans*	subsp. silvestris	(Villiers) Brenan	n/a	Albersia	Eurasia	
Amaranthus	*graecizans*	subsp. thellugianus	(Nevski) Gusev	n/a	Albersia	Eurasia	

Genus	Species	Subspecies	Authority	Subgenus Mosyakin and Robertson	Subgenus Bayón	Geographical origins: Continent(s)	Geographical origins: Areas within continent(s)
Amaranthus	*greggii*		S. Watson	Acnida sect. Saueranthus	n/a	North America	Coastal Louisiana, Texas, Mexico
Amaranthus	*hybridus*		L.	Amaranthus sect. Amaranthus	Amaranthus	North America	Eastern North America
Amaranthus	*hypochondriacus*		L.	Amaranthus sect. Amaranthus	Amaranthus	North America	Southwest Mexico
Amaranthus	*induratus*		C.A. Gardner ex J. Palmer & Mowatt	Albersia sect. Pentamorion?	Albersia	Australia	Northern and Western Australia
Amaranthus	*interruptus*		R. Br.	Albersia sect. Pentamorion	Albersia	Australia	Northern, Northwestern, and Central Australia
Amaranthus	*kloosianus*		Hunz.	Albersia sect. Pentamorion	Albersia	South America	Argentina (Jujuy, La Rioja, Salta)
Amaranthus	*looseri*		Suess.	Albersia sect. Pentamorion?	Albersia	South America	Chile
Amaranthus	*macrocarpus*		Benth.	Albersia sect. Pentamorion	Albersia	Australia	Eastern Australia
Amaranthus	*mitchellii*		Benth.	Albersia sect. Pentamorion	Albersia	Australia	Central and western Australia
Amaranthus	*muricatus*		(Moq.) Hieronymus	Albersia sect. Pentamorion	Albersia	South America	Argentina, Bolivia, Paraguay, Uruguay
Amaranthus	*palmeri*		S. Watson	Acnida sect. Saueranthus	n/a	North America	Southwest U.S. and northern MX
Amaranthus	*persimilis*		Hunz.	Albersia sect. Pentamorion	Albersia	South America	Argentina (Catamarca, Mendoza, San Juan, Tucumán)
Amaranthus	*polygonoides*		L.	Albersia sect. Pentamorion?	Albersia	North America/Central America/South America	US Gulf Coast, West Indies, northern South America
Amaranthus	*powellii*	subsp. bouchonii	(Thell.) Costea & Carretero	Amaranthus sect. Amaranthus (as A. bouchonii)	n/a	North America	Southwest U.S. and northern Mexico
Amaranthus	*powellii*	subsp. powellii	S. Watson	Amaranthus sect. Amaranthus	Amaranthus	North America	
Amaranthus	*praetermissus*		Brenan	Albersia sect. Pyxidium?	Albersia	Africa	Southern Africa
Amaranthus	*pumilus*		Raf.	Albersia sect. Pentamorion?	Albersia	North America	US Atlantic coast

Genus	Species	Subspecies	Authority	Subgenus Mosyakin and Robertson	Subgenus Bayón	Geographical origins: Continent(s)	Geographical origins: Areas within continent(s)
Amaranthus	*quitensis*		Kunth	Amaranthus sect. Amaranthus	Amaranthus (as subspecies of *A. hybridus*)	South American	Andean highlands
Amaranthus	*retroflexus*		L.	Amaranthus sect. Amaranthus	Amaranthus	North America	Central and eastern North America
Amaranthus	*rhombeus*		R. Br.	Albersia sect. Pyxidium?	Albersia	Australia	Coast of Northern Territory and Queensland
Amaranthus	*scariosus*		Benth.	Albersia sect. Pyxidium?	Amaranthus	North America/Central America	West coast of Mexico and Central America
Amaranthus	*sclerantoides*		(Andersson) Andersson	Albersia sect. Pyxidium?	Albersia	South America	Galápagos Islands
Amaranthus	*spinosus*		L.	Amaranthus sect. Centrusa	Amaranthus	North America/Central America/South America	Neotropics
Amaranthus	*squamulatus*		(Andersson) B.L. Rob.	Albersia sect. Pentamorion?	Albersia	Ecuador	Galápagos Islands and coastal Ecuador
Amaranthus	*standleyanus*		Parodi ex Covas	Albersia sect. Pentamorion	Albersia	South America	Central and northwest Argentina and Paraguay
Amaranthus	*tamaulipensis*		Henrickson	Albersia sect. Pyxidium?	Albersia	North America	Texas, northern Mexico
Amaranthus	*thunbergii*		Moq.	Albersia sect. Pyxidium	Albersia	Africa	Southern Africa
Amaranthus	*tricolor*		L.	Albersia sect. Pyxidium	Albersia	Eurasia	Tropical Asia
Amaranthus	*tuberculatus*		(Moq.) J.D. Sauer	Acnida sect. Acnida	n/a	North America	Midwest US
Amaranthus	*x tucsonensis*		Henrickson	n/a	n/a	North America	Southwest US
Amaranthus	*undulatus*		R. Br.	Albersia sect. Pentamorion	Albersia	Australia	Northern and northwestern Australia and Queensland

Genus	Species	Subspecies	Authority	Subgenus Mosyakin and Robertson	Subgenus Bayón	Geographical origins: Continent(s)	Geographical origins: Areas within continent(s)
Amaranthus	*urceolatus*		Benth.	Albersia sect. Pentamorion?	Albersia	South America	Northwest Argentina, Peru, and Ecuador
Amaranthus	*viridis*		L.	Albersia sect. Blitopsis	Albersia	South America	South American tropics
Amaranthus	*vulgatissimus*		Speg.	Albersia sect. Pentamorion	Albersia	South America	Argentina
Amaranthus	*watsonii*		Standley	Acnida sect. Saueranthus	n/a	North America	Arizona, California, Baja California, Sonora
Amaranthus	*wrightii*		S. Watson	Amaranthus sect. Amaranthus?	Amaranthus	North America	Southwest US
Chamissoa	*altissima*		(Jacq.) Kunth	n/a	n/a	North America/Central America/South America	Mexico to Brazil
Pleuropterantha	*revoilii*		Franch.	n/a	n/a	Africa	Ethiopia, Somalia

Table 1.2. Specimens sampled for the molecular phylogeny, including the abbreviations for the specimens used in the phylogenetic tree figures. If material or seeds was obtained from the USDA GRIN Database (Agricultural Research Service in Ames, IA), a PI number or Ames number is listed. If leaf tissue was obtained from another source, the collector and collection number, as well as the herbarium and herbarium accession number if available, are listed instead.

Genus	Species	Subspecies	Abbreviation of name used in phylogenetic trees	USDA PI Number	Collector	Collection Number	Herbarium and Herbarium Accession Number	Originally collected from: Country: State/Province
Amaranthus	*acanthochiton*		Acantho	PI 632238				US: Texas
Amaranthus	*acutilobus*		Acutilobus	PI 633579				Germany
Amaranthus	*albus*		AlbusCan	PI 633580				Canada: Saskatchewan
Amaranthus	*albus*		AlbusSA		Le Roux sub Boatwright	508	NBG	South Africa
Amaranthus	*anderssonii*		Andersson		H. Jäger		CDF 13607	Ecuador: Galápagos Islands
Amaranthus	*arenicola*		Arenicola	PI 607459				US: Kansas
Amaranthus	*australis*		Australis	PI 553076				US: Florida
Amaranthus	*australis*		AustralJRA		J. Richard Abbott	25276	FLAS 232341	US: Florida
Amaranthus	*blitoides*		Blitoides	PI 553079				US: Iowa
Amaranthus	*blitoides*		BlitoidesNM	Ames 27956				US: New Mexico
Amaranthus	*blitum*	subsp. blitum	BlitumB	PI 606751				Switzerland
Amaranthus	*blitum*	subsp. emarginatum	BlitumE		J. Richard Abbott	24900	FLAS 226902	US: Florida
Amaranthus	*blitum*	subsp. oleraceus	BlitumO	PI 606282				Bangladesh
Amaranthus	*blitum*	subsp. pseudogracilis	BlitumP	PI 632245				US: North Carolina
Amaranthus	*californicus*		Californicus	PI 595319				US: California
Amaranthus	*cannabinus*		Cannabinus	PI 568124				US: Virginia
Amaranthus	*cardenasianus*		Cardenas		D. Rocabado et al.	499	MO 4787435	Bolivia
Amaranthus	*caudatus*		CaudatusARG	Ames 15178				Argentina
Amaranthus	*caudatus*		CaudatusIND	PI 166045				India
Amaranthus	*centralis*		Centralis		D.E. Albrecht	8892	CANB 527441	Australia: Northern Territory
Amaranthus	*clementii*		ClemCran		R. Cranfield	9595	CANB 496410	Australia: Western Australia
Amaranthus	*clementii*		ClemCress		I.D. Cresswell	97V1-OP-03	CANB 497238	Australia: Western Australia
Amaranthus	*cochleitepalus*		Cochleitep		D.E. Albrecht	9153	CANB 577421	Australia: Northern Territory
Amaranthus	*crassipes*		Crassipes	PI 642743				US: Texas

Genus	Species	Subspecies	Abbreviation of name used in phylogenetic trees	USDA PI Number	Collector	Collection Number	Herbarium and Herbarium Accession Number	Originally collected from: Country: State/Province
Amaranthus	*crassipes*		CrassipesTX2	PI 649302				US: Texas
Amaranthus	*crispus*		Crispus	PI 633582				Hungary
Amaranthus	*cruentus*		CruentusIND	PI 566897				India
Amaranthus	*cruentus*		CruentusMX	PI 477913				Mexico
Amaranthus	*aff. cuspidifolius*	(affinity species)	AffCuspid		R. Bates	50387	CANB 689602	Australia: South Australia
Amaranthus	*cuspidifolius*		Cuspid605		J. Palmer	605	CANB 599739	Australia: Western Australia
Amaranthus	*cuspidifolius*		Cuspid699		J. Palmer	699	CANB 775595	Australia: Western Australia
Amaranthus	*deflexus*		DeflexARG	Ames 15314				Argentina
Amaranthus	*deflexus*		DeflexC		S. Torres Robles	400		Argentina: Buenos Aires
Amaranthus	*deflexus*		DeflexPort	PI 633576				Portugal
Amaranthus	*dubius*		DubiusC	PI 642739				Cuba
Amaranthus	*dubius*		DubiusVZ	Ames 15320				Venezuela
Amaranthus	*fimbriatus*		Fimbriat612	PI 612855				US: Arizona
Amaranthus	*fimbriatus*		Fimbriat662	PI 662285				US: Arizona
Amaranthus	*floridanus*		Floridanus	PI 553078				US: Florida
Amaranthus	*graecizans*	subsp. aschersonianus	GraecAsch	PI 288277				India
Amaranthus	*graecizans*	subsp. silvestris	GraecSilv	PI 658732				Portugal
Amaranthus	*graecizans*	subsp. silvestris	GraecSilvUS	PI 604196				Ecuador
Amaranthus	*graecizans*	subsp. thellugianus	GraecThell	PI 549157				Mauritania
Amaranthus	*greggii*		GreggiiLA	PI 667170				US: Louisiana
Amaranthus	*greggii*		GreggiiTX	PI 632240				US: Texas
Amaranthus	*hybridus*		HybridusCOR2		K. Waselkov			US: Missouri
Amaranthus	*hybridus*		HybridusGuat	Ames 21999				Guatemala
Amaranthus	*hybridus*		HybridusSpE2		K. Waselkov			US: Missouri
Amaranthus	*hypochondriacus*		HypochonIND	PI 477915				India
Amaranthus	*hypochondriacus*		HypochonMX	PI 477917				Mexico
Amaranthus	*induratus*		Induratus		A.A. Mitchell	5749	CANB 556042	Australia: Western Australia
Amaranthus	*interruptus*		Interruptus		L.A. Craven et al.	9659	CANB 498997	Australia: Northern Territory
Amaranthus	*kloosianus*		Kloosianus		A. Plos and P. Simon	133		Argentina: Tucumán

Genus	Species	Subspecies	Abbreviation of name used in phylogenetic trees	USDA PI Number	Collector	Collection Number	Herbarium and Herbarium Accession Number	Originally collected from: Country: State/Province
Amaranthus	*looseri*		Looseri		M. Muñoz	5103	SGO	Chile
Amaranthus	*macrocarpus*		Macrocarpus		J. Hosking	3238		Australia
Amaranthus	*mitchellii*		Mitchellii		A.A. Mitchell	8726B	CANB 711440	Australia: Western Australia
Amaranthus	*muricatus*		MuricatusC		J. Hurrell et al.	3881		Argentina: Buenos Aires
Amaranthus	*muricatus*		MuricatusPS		A. Plos and P. Simon	158		Argentina: Salta
Amaranthus	*mystery species*		MysterySps		F. Zuloaga	12119		Argentina
Amaranthus	*new species*		NewSps		K. Waselkov	222	CDF	Ecuador: Galápagos Islands
Amaranthus	*palmeri*		Palmeri	PI 632235				US: Arizona
Amaranthus	*palmeri*		PalmeriAZ2	PI 612856				US: Arizona
Amaranthus	*palmeri*		PalmeriMX	PI 633593				Mexico
Amaranthus	*persimilis*		Persimilis		A. Plos and P. Simon	135		Argentina
Amaranthus	*polygonoides*		Polygon	PI 658733				US: Texas
Amaranthus	*powellii*	subsp. bouchonii	PowelliiB	PI 572261				Germany
Amaranthus	*powellii*	subsp. powellii	PowelliiP	PI 604671				US: Washington
Amaranthus	*powellii*		PowelliiZ		F. Zuloaga	11496		Argentina: Jujuy
Amaranthus	*praetermissus*		Praetermiss		J. Manning			South Africa
Amaranthus	*pumilus*		PumilusNC	PI 553083				US: North Carolina
Amaranthus	*pumilus*		PumilusNJ		Mt. Cuba Center	2000211*A		US: Delaware
Amaranthus	*pumilus*		PumilusSC	PI 553085				US: South Carolina
Amaranthus	*quitensis*		Quitensis	PI 511745				Ecuador
Amaranthus	*retroflexus*		Retroflexus	PI 603852				US: Iowa
Amaranthus	*rhombeus*		RhombeusG		B. Gray	7948	CANB 670451	Australia: Queensland
Amaranthus	*rhombeus*		RhombeusR		A.P. Roberts et al.	804	CANB 693250	Australia: Northern Territory
Amaranthus	*scariosus*		Scariosus		I. Coronado G. and R.M. Rueda	3570	MO 6180339	Nicaragua
Amaranthus	*sclerantoides*		SclerantSC		K. Waselkov	206		Ecuador: Galápagos Islands
Amaranthus	*spinosus*		SpinosusNC	PI 632248				US: North Carolina
Amaranthus	*spinosus*		SpinosusS		B. Summers	6179	MO	US: Missouri
Amaranthus	*squamulatus*		SquamulSC		K. Waselkov	205		Ecuador: Galápagos Islands

Genus	Species	Subspecies	Abbreviation of name used in phylogenetic trees	USDA PI Number	Collector	Collection Number	Herbarium and Herbarium Accession Number	Originally collected from: Country: State/Province
Amaranthus	*squamulatus*		SquamulST		K. Waselkov	204		Ecuador: Galápagos Islands
Amaranthus	*standleyanus*		StandleyPS		A. Plos and P. Simon	132		Argentina: Tucumán
Amaranthus	*standleyanus*		StandleyZ		F. Zuloaga	11559		Argentina
Amaranthus	*tamaulipensis*		Tamaulip	PI 642738				Cuba
Amaranthus	*thunbergii*		Thunberg1889		HK	871	NPGRC 1889	Namibia
Amaranthus	*thunbergii*		Thunberg2111		HK	1038	NPGRC 2111	Namibia
Amaranthus	*tricolor*		TricolorMP	PI 599683				India: Madhya Pradesh
Amaranthus	*tricolor*		TricolorTN	PI 566899				India: Tamil Nadu
Amaranthus	*tuberculatus*		TuberculCHE2		K. Waselkov			US: Kansas
Amaranthus	*tuberculatus*		TuberculPEK2		K. Waselkov			US: Illinois
Amaranthus	*x tucsonensis*		Tucsonen	Ames 30697				US: Arizona
Amaranthus	*undulatus*		Undulat580		J. Palmer	580	CANB 599392	Australia: Western Australia
Amaranthus	*undulatus*		Undulat652		J. Palmer	652	CANB 686336	Australia: Western Australia
Amaranthus	*urceolatus*		Urceolatus		S. Llatas Quiroz	3057	MO 3318704	Peru
Amaranthus	*viridis*		ViridisBZ	PI 652434				Brazil
Amaranthus	*viridis*		ViridisJ	PI 540445				Indonesia: Java
Amaranthus	*vulgatissimus*		VulgatC		J.A. Tolaba and R. Alacón	3427		Argentina: Salta
Amaranthus	*vulgatissimus*		VulgatPS		A. Plos and P. Simon	108		Argentina: Tucumán
Amaranthus	*watsonii*		Watsonii		A.C. Sanders et al.	8768	MO 4919874	Mexico: Sonora
Amaranthus	*wrightii*		Wrightii242	PI 632242				US: Texas
Amaranthus	*wrightii*		WrightiiTX2	PI 632243				US: Texas
Chamissoa	*altissima*		Chamissoa		L. Alvarado-Cárdenas et al.	1182	MO 6327402	Mexico: Chiapas
Chamissoa	*altissima*		Chamissoa3		Carrasco et al.	272	MO 4821722	Bolivia: Santa Cruz
Pleuropterantha	*revoilii*		Pleuropter		M. Thulin	10831	UPS	Somalia

Table 1.3. Genetic regions used in the phylogeny, with primers, source of primers, length of the aligned sequence, and number and percent of variable sites and parsimony-informative sites, with and without the outgroups.

					With Outgroups		*Amaranthus* Alone	
Region	**Primer Name**	**Primers (5'-3')**	**Source**	**Aligned basepairs**	**#/% variable sites**	**#/% parsimony-informative sites**	**#/% variable sites**	**#/% parsimony-informative sites**
A36	A36F	TGGTTATCCGTGCCTTTCTC	Lawton-Rauh lab	734	145 (20%)	93 (13%)	81 (11%)	58 (8%)
	A36R	CAGGACCTGGATTCTTTCCA	Lawton-Rauh lab					
	A361F	GCAACCTGTGCCACAGGACCTG	internal					
	A361R	CAGGTCCTGTGGCACAGGTTGC	internal					
G3PDH	G3F	AGGGTCTCATGACAACTGTTCACTCT	redesigned from Strand et al., 1997	884	272 (31%)	170 (19%)	245 (28%)	168 (19%)
	G3R2	TCACCAACGAAGTCGGTGGAA	redesigned from Strand et al., 1997					
	G3BIF	CACTGGAGCAGCCAAGGTAT	internal					
ITS	ITS4	TCCTCCGCTTATTGATATGC	Lawton-Rauh lab	680	184 (27%)	147 (22%)	113 (17%)	80 (12%)
	ITS5	GGAAGTAAAAGTCGTAACAAGG	Lawton-Rauh lab					
Waxy	WXF12	GGTCTTGGTGATGTCCTTGG	designed from Park et al., 2010	1252	428 (34%)	252 (20%)	343 (27%)	233 (19%)
	WXR7	AGGCAAATCTTCCTTGATATACAATA	designed from Park et al., 2010					
	WXF5	TAATATGTGCTTCAGGCAGCT	internal					
	WXR5	GAAGTTCGGATTGTTGTTGAGA	internal					
matK/trnK	TrnKF1	ATCATGGGGTTGCTAACTCA	Muller and Borsch, 2005	2400	358 (15%)	224 (9%)	184 (8%)	119 (5%)
	TrnKR1	AACTAGTCGGATGGAGTAG	Muller and Borsch, 2005					
	TrnKR31	GGCATCTTTCAACCAATAGCGAAGAG	internal					
	MatKF	CGATCTATTCATTCAATATTTC	Lawton-Rauh lab					
	MatKR	TCTAGCACACGAAAGTCGAAGT	Lawton-Rauh lab					
	MatK1F	AAGAACCTTTTCTGCATTATGTTCGG	internal					
trnC-trnD	trnL_C	CGAAATCGGTAGACGCTACG	Shaw et al., 2005	634	70 (11%)	43 (7%)	40 (6%)	22 (3%)
	trnL_D	GGGGATAGAGGGACTTGAAC	Shaw et al., 2005					
Nuclear concatenated		n/a	n/a	3549	965 (27%)	610 (17%)	703 (20%)	491 (14%)
Chloroplast concatenated		n/a	n/a	3034	428 (14%)	267 (9%)	224 (7%)	141 (5%)

Table 1.4. Molecular models of evolution chosen for each dataset as the best fit by MrModeltest 2.3. Two different criteria are available from MrModelTest2: the hierarchical likelihood ratio test (hLRT) and the Akaike information criterion (AIC).

Dataset	hLRT	AIC
A36	GTR+Γ	GTR+Γ
G3PDH	GTR+Γ	GTR+Γ
ITS	GTR+I+Γ	GTR+I+Γ
Waxy	HKY+Γ	HKY+Γ
Concatenated chloroplast	GTR+Γ	GTR+I+Γ

Table 1.5a. Values for 12 morphological and ecological traits (5 quantitative, 7 qualitative) for the sampled species of *Amaranthus*. Literature sources for the trait values are mentioned in the text, and a key to the qualitative trait values is given in Table 1.5b.

Species of *Amaranthus*	Agricultural weed rank	Agricultural weed status	Problem-atic weed status	Mating system	Maximum plant size (cm)	Maximum leaf length (cm)	Utricle dehisce-nce	Average seed diameter (mm)	Habitat 1 (Beaches)	Habitat 2 (Ruderal)	Habitat 3 (Water)	Habitat 4 (Natural disturbance)	Number of GBIF cells occupied	Range expanded by human s	Maximu m Elevation (m)
A. acanthochiton	1	1	1	2	80	8	1	1.15	2	1	2	1	13	1	2000
A. acutilobus	2	2	1	1	35	1.8	1	1.15	1	2	1	2	4	2	2600
A. albus	3	2	2	1	100	1.5	2	0.8	1	2	2	1	500	2	2200
A. anderssonii	1	1	1	1	30	1	1	0.9	2	1	2	1	3	1	20
A. arenicola	2	2	1	2	200	8	2	1.75	1	1	2	1	141	2	2000
A. australis	1	1	1	2	900	20	1	1.1	1	1	2	2	54	1	100
A. blitoides	2	2	1	1	100	4	2	1.45	1	2	2	1	500	2	2200
A. blitum	2	2	1	1	60	6	1	1.4	1	2	1	1	312	2	1000
A. californicus	1	1	1	1	50	3	2	0.85	1	2	2	1	72	1	2800
A. cannabinus	1	1	1	2	300	20	1	1	1	1	2	2	41	1	50
A. cardenasianus	1	1	1	1	70	6.5	2	1	1	1	1	2	2	1	2000
A. caudatus	1	1	1	1	250	20	2	1.1	1	2	1	2	217	2	3000
A. centralis	1	1	1	1	60	5.5	1	1.3	1	1	2	1		1	500
A. clementii	1	1	1	1	30	5	2	1.38	1	1	1	1	11	1	500
A. cochleitepalus	1	1	1	1	20	1.5	1	0.8	1	1	2	2	35	1	500
A. crassipes	1	1	1	1	60	4.5	1	1.2	2	2	2	1	47	2	1300
A. crispus	1	1	1	1	50	2.5	1	0.85	1	2	1	2	40	2	500
A. cruentus	1	1	1	1	200	20	2	1.4	1	2	1	2	235	2	
A. cuspidifolius	1	1	1	1	30	4	1	1.25	1	2	2	1	56	1	500
A. deflexus	2	2	1	1	50	8	1	1.1	1	2	1	1	261	2	500
A. dubius	1	1	1	1	100	12	2	0.9	1	2	2	1	140	2	1000
A. fimbriatus	1	1	1	1	100	10	2	0.9	1	2	1	1	73	1	1700
A. floridanus	1	1	1	2	150	20	1	0.85	2	1	2	1	6	1	10
A. graecizans	1	1	1	1	90	5	2	1.15	1	2	1	1	279	2	
A. greggii	1	1	1	2	100	4	1	1.45	2	1	2	1	29	1	50
A. hybridus	3	2	2	1	250	15	2	1.15	1	2	1	1	600	2	2500
A. hypochondriacus	1	1	1	1	250	12	2	1.2	1	2	1	2	131	2	

Species	Agricultural weed rank	Agricultural weed status	Problem-atic weed status	Mating system	Maximum plant size (cm)	Maximum leaf length (cm)	Utricle dehisce-nce	Average seed diameter (mm)	Habitat 1 (Beaches)	Habitat 2 (Ruderal)	Habitat 3 (Water)	Habitat 4 (Natural disturbance)	Number of GBIF cells occupied	Range expanded by human s	Maximu m Elevation (m)
A. induratus	1	1	1	1	90	7	1	1.4	1	2	2	1		1	500
A. interruptus	1	1	1	1	60	5	1	1	1	1	2	1	85	2	500
A. kloosianus	1	1	1	1	60	3.8	1	1	1	2	1	1	3	1	2500
A. looseri	1	1	1	1	6	0.8	1	1	1	1	2	2	1	1	500
A. macrocarpus	2	2	1	1	60	2.5	1	1.13	1	2	2	1	61	2	500
A. mitchellii	2	2	1	1	50	3.5	1	1.4	1	2	2	1	105	2	500
A. muricatus	1	1	1	1	40	8	1	1.1	1	2	1	1	90	2	1000
A. palmeri	3	2	2	2	300	7	2	1.1	1	2	2	1	227	2	1000
A. persimilis	1	1	1	1	100	6.5	1	1.25	1	2	1	1	2	1	2000
A. polygonoides	1	1	1	1	50	4	1	0.9	2	2	2	1	40	2	500
A. powellii	3	2	2	1	200	8	2	1.2	1	2	2	1	370	2	2500
A. praetermissus	1	1	1	1	100	4	2	1.1	1	1	1	2	20	1	1000
A. pumilus	1	1	1	1	50	1.5	1	2.5	2	1	2	1	15	1	10
A. retroflexus	3	2	2	1	200	15	2	1.15	1	2	2	1	500	2	2500
A. rhombeus	1	1	1	1	22	3	2	1	2	1	2	1	9	1	50
A. scariosus	1	1	1	1	250	6	2	0.9	1	2	1	1	17	1	500
A. sclerantoides	1	1	1	1	40	2.5	1 or 2	1	2	1	2	1		1	10
A. spinosus	3	2	2	1	200	15	1	0.85	1	2	1	1	500	2	700
A. squamulatus	1	1	1	1	100	6	1	1.07	2	1	1	1	3	1	150
A. standleyanus	2	2	1	1	50	8	1	1.15	1	2	1	1	54	2	1500
A. tamaulipensis	1	1	1	1	60	2.7	2	1.1	1	2	1	1	1	1	100
A. thunbergii	2	2	1	1	100	2	2	1.2	1	2	2	1	84	1	1400
A. tricolor	1	1	1	1	150	12	2	1	1	2	1	2	37	2	
A. tuberculatus	3	2	2	2	300	15	1 or 2	0.85	1	2	2	1	241	2	1000
A. undulatus	1	1	1	1	100	4.5	2	1.15	2	1	2	1	66	1	500
A. urceolatus	1	1	1	1	80	4	1	0.8	1	1	1	2	5	1	3600
A. viridis	3	2	2	1	50	7	1	1.25	1	2	1	1	461	2	1000
A. vulgatissimus	1	1	1	1	30	3	1	1.3	1	1	1	2	4	1	500
A. watsonii	1	1	1	2	100	8	2	1.1	2	1	1	1	27	1	100
A. wrightii	1	1	1	1	100	6	2	1	1	1	2	1	14	1	2000

Table 1.5b. Key for Table 1.5a, explaining the meaning of the values for each qualitative character.

Agricultural weed rank	Scales of weediness: 1 = not a weed, 1 = occasionally found in crop fields and/or "casual" agricultural weeds, 3 = frequently associated with agriculture and problematic agriculturally
Agricultural weed status	Absolute weediness 1: 1 = never a weed, 2 = sometimes or always a weed
Problematic weed status	Absolute weediness 2: 1 = not a problematic weed, 2 = a problematic weed
Mating system	1 = monoecious, 2 = dioecious
Maximum plant size	quantitative
Maximum leaf length	quantitative
Utricle dehiscence	1 = indehiscent, 2 = dehiscent
Average seed diameter	quantitative
Habitat 1	1 = does not grow on beaches, 2 = grows on beaches
Habitat 2	1 = not a ruderal weed, 2 =a ruderal weed (roadsides, railroads, pastures, other anthropogenically disturbed areas)
Habitat 3	1 = not associated with water, 2 = associated with water (riverbanks, streams, wet places)
Habitat 4	1 = in naturally disturbed areas, 2 = not in naturally disturbed areas
Number of GBIF cells occupied	quantitative
Expanded geographical range due to human seed movement	1 = no, 2 = yes
Maximum elevation	quantitative

Table 1.6. Results of statistical tests for association of 13 morphological and ecological traits with 3 different agricultural invasiveness metrics. df = degrees of freedom. Red type denotes significant rejection of the null hypothesis of no association between a variable and invasiveness. Ranking of groups shows the results of the Bonferroni post-hoc tests.

Qualitative traits: Chi-square tests

Trait:	Mating system		Utricle dehiscence		Habitat 1 (Beaches)		Habitat 2 (Ruderal)		Habitat 3 (Water)		Habitat 4 (Natural disturbance)		Range expanded by humans	
Metric	**Chi-square value (df)**	**Result**	**Chi-square value (df)**	**Result**	**Chi-square value (df)**	**Result**	**Chi-square value (df)**	**Result**	**Chi-square value (df)**	**Result**	**Chi-square value (df)**	**Result**	**Chi-square value (df)**	**Result**
Agricaltural weed rank	0.6773 (df=2)	Can't reject null	2.2018 (df=2)	Can't reject null	6.4600 (df=2)	Reject null	11.7029 (df=2)	Reject null	0.2746 (df=2)	Can't reject null	4.8439 (df=2)	Can't reject null	21.4459 (df=2)	Reject null
Agricultural weed status	0.0629 (df=1)	Can't reject null	0.1874 (df=1)	Can't reject null	6.4600 (df=1)	Reject null	11.6199 (df=1)	Reject null	0.1923 (df=1)	Can't reject null	4.5617 (df=1)	Reject null	21.2362 (df=1)	Reject null
Problematic weed status	0.5938 (df=1)	Can't reject null	1.8697 (df=1)	Can't reject null	2.4816 (df=1)	Can't reject null	6.5051 (df=1)	Reject null	0.2470 (df=1)	Can't reject null	3.0299 (df=1)	Can't reject null	10.3401 (df=1)	Reject null

Quantitative traits: Independent sample t-tests or ANOVAs

Trait:	Log maximum plant size			Log maximum leaf length			Squareroot average seed diameter			Log number of GBIF cells occupied			Squareroot maximum elevation		
Metric	**F (df)**	**P-value**	**Ranking of groups**	**F (df)**	**P-value**	**Ranking of groups**	**F (df)**	**P-value**	**Ranking of groups**	**F (df)**	**P-value**	**Ranking of groups**	**F (df)**	**P-value**	**Ranking of groups**
Agricaltural weed rank	4.058 (df=2)	0.023	3 > 1 and 2	1.976 (df=2)	0.149	n/a	6.062 (df=2)	0.004	2 > 1 and 3	15.944 (df=2)	<0.001	3 > 2 > 1	3.947 (df=2)	0.026	3 > 2 > 1
	t (df)	**P-value**		**t (df)**	**P-value**		**t (df)**	**P-value**		**t (df)**	**P-value**		**t (df)**	**P-value**	
Agricultural weed status	-1.427 (df=55)	0.159		-0.461 (df=55)	0.646		-1.830 (df=54)	0.073		-5.087 (df=52)	<0.001		-2.756 (df=1)	0.008	
Problematic weed status	-2.852 (df=55)	0.006		-1.813 (df=55)	0.075		-0.975 (df=54)	0.334		-9.780 (df=48.9)	<0.001		-2.128 (df=51)	0.038	

Table 1.7. Species of *Amaranthus* that were not sampled in this study, listed with their geographical range and morphological affinities from previously published literature. Species marked with an asterisk have not been verified as distinct taxonomic units by Néstor Bayón and may be synonymous with other monoecious species.

Species	Geographical Range	Morphological affinities	Source for morphological affinities
A. acanthobracteatus	Northern Mexico	Sister to *A. acanthochiton*	Henrickson, 2004
A. asplundii	Ecuador to Chile/Argentina	Subgenus Amaranthus	Bayón, in review
*A. brandegei**	SW U.S. and adjacent Mexico	Similar to *A. torreyi*	Mosyakin and Robertson, 2003
A. brownii	Hawaii	?	
A. capensis	South Africa	Similar to *A. dinteri* and *A. thunbergii*	Brenan, 1981
*A. celosioides**	South America	Subgenus Amaranthus?	Bayón, in review
*A. congestus**	Venezuela	?	
A. dinteri	South Africa	Similar to *A. capensis* and *A. thunbergii*	Brenan, 1981
A. furcatus	Galápagos Islands	Similar to *A. sclerantoides* (dubiously distinct)	Eliasson, 1985
A. hunzikeri	Northwest Argentina	Similar to *A. kloosianus*	Bayón, in review
A. lombardoi	Uruguay	Similar to *A. viridis* and *A. deflexus*	Bayón, in review
A. minimus	Cuba	?	
A. obcordatus	Arizona and adjacent Mexico	Similar to *A. fimbriatus*	Bayón, in review
A. pedersenii	Argentina	Similar to *A. kloosianus*	Bayón, in review
A. peruvianus	Peru, Bolivia, and Argentina	Similar to *A. looseri*?	Bayón, in review
A. rosengurtii	Uruguay and Argentina	Similar to *A. muricatus*	Hunziker, 1966
A. schinzianus	South Africa	Similar to *A. praetermissus*	Brenan, 1981
A. scleropoides	Texas and adjacent Mexico	Similar to *A. crassipes*	Mosyakin and Robertson, 2003; Bayón, in review
*A. sparghaniocephalus**	Ethiopia	?	
*A. tenuifolius**	Pakistan	?	
A. torreyi	SW U.S. and adjacent Mexico	Similar to *A. fimbriatus*	Bayón, in review
A. viscidulus	New Mexico	?	

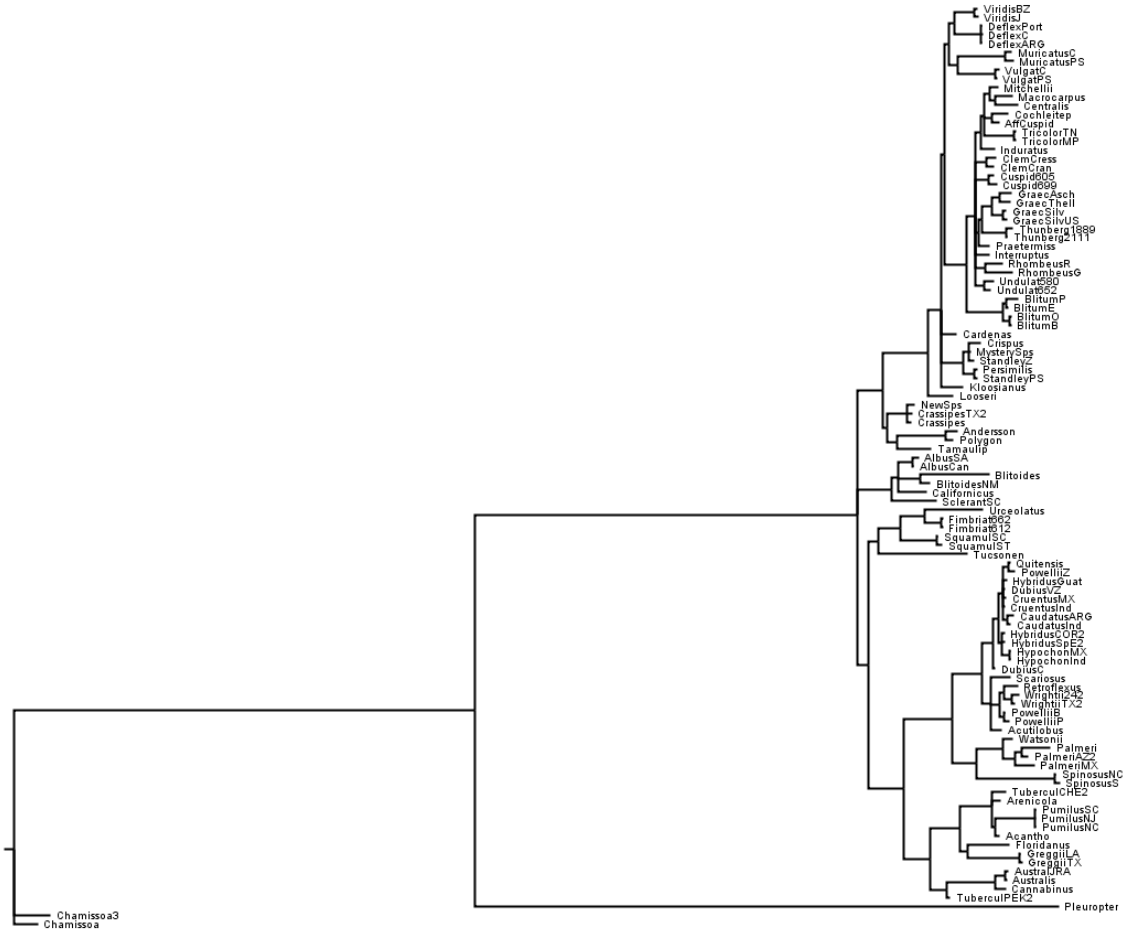

Figure 1.1. Bayesian 50% majority-rule consensus tree with branch lengths (and without posterior probability values) for the partitioned model for the concatenated nuclear dataset.

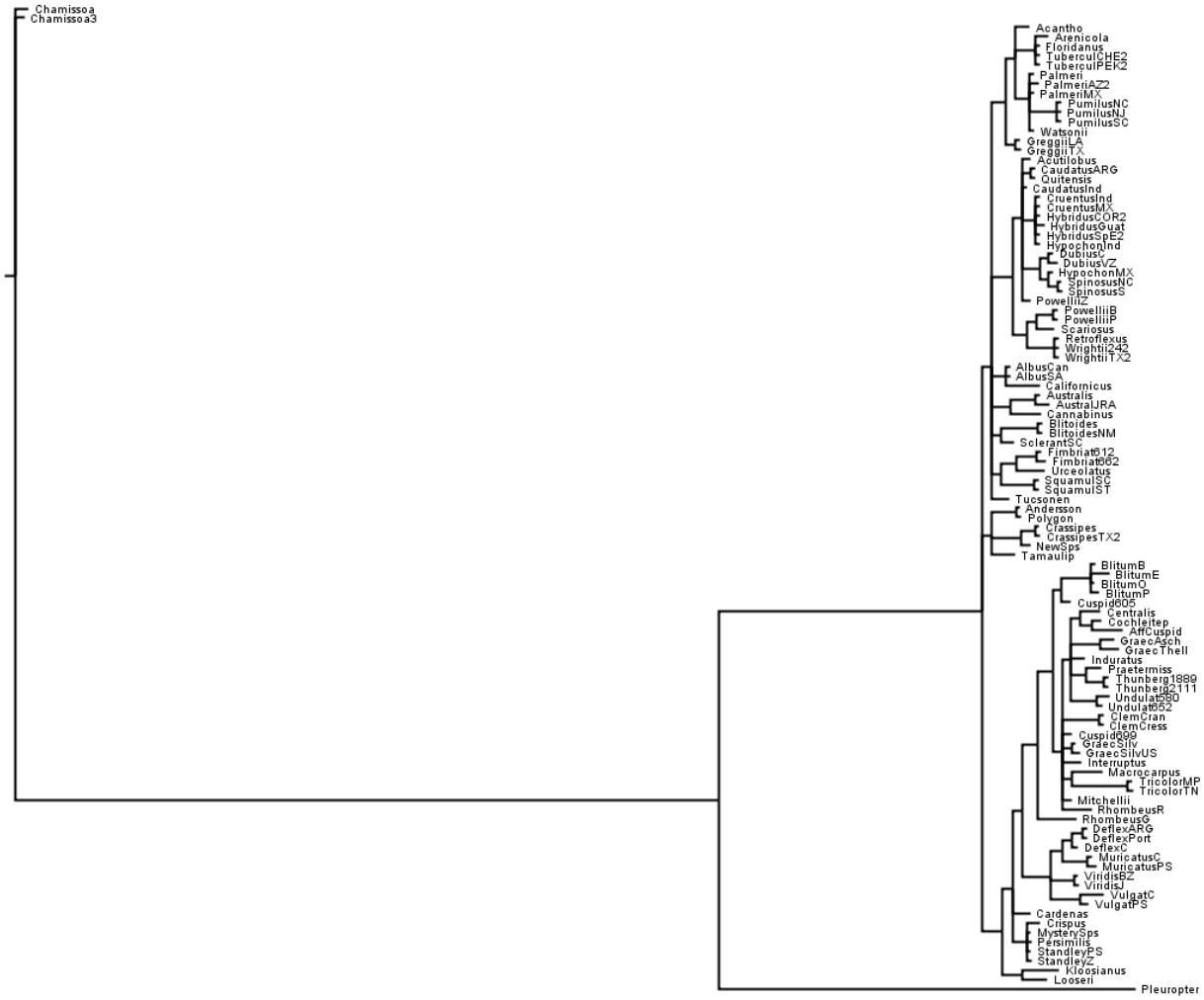

Figure 1.2. Bayesian 50% majority-rule consensus tree with branch lengths (and without posterior probability values) for the GTR+Γ model for the concatenated chloroplast dataset.

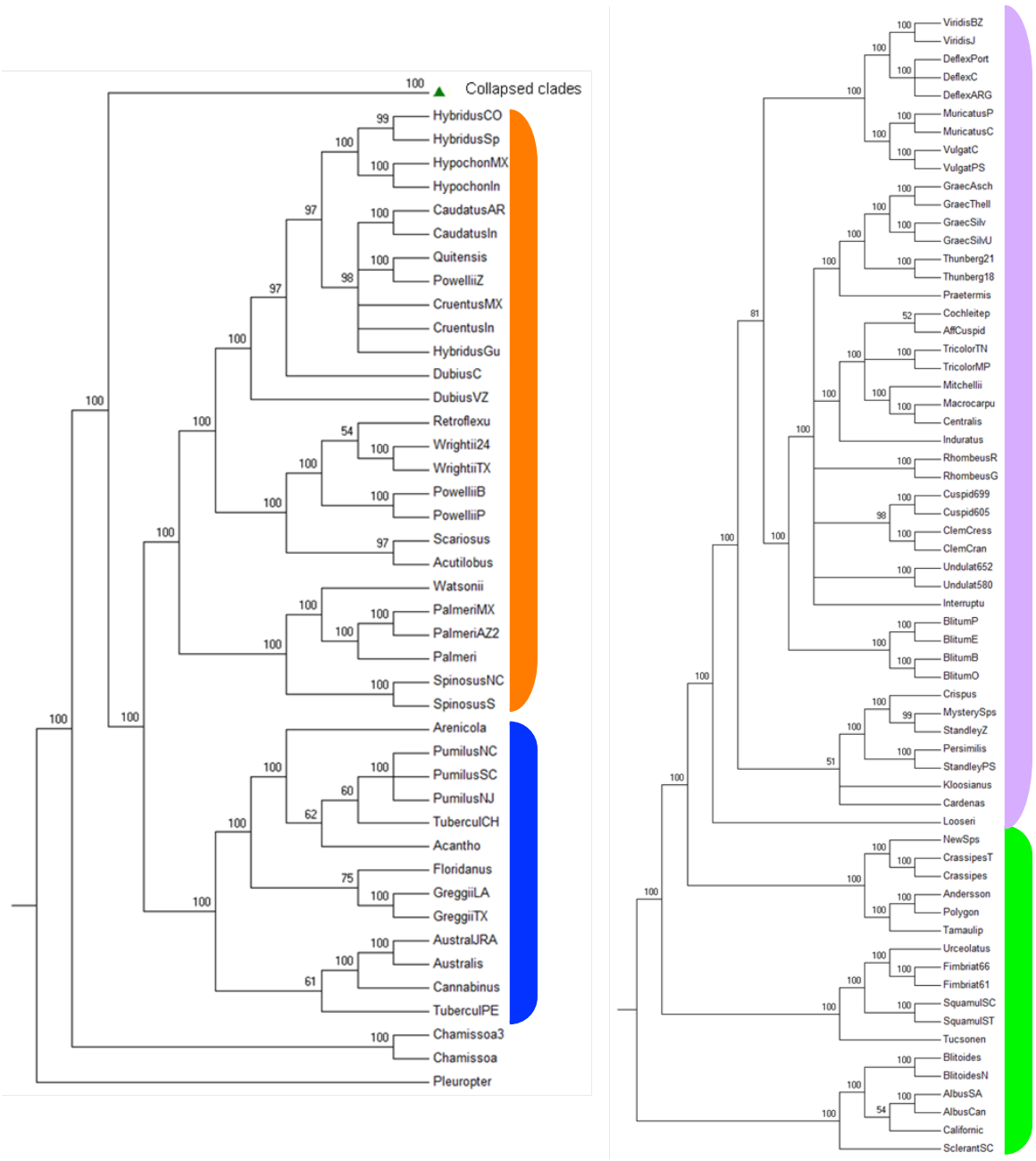

Figure 1.3. Maximum parsimony 50% majority-rule consensus tree of the most-parsimonious trees for the concatenated nuclear dataset. The right-hand side shows the ESA + South American and Galápagos clades that have been collapsed in the left-hand tree. The major clades are shown with colored bars: Purple = the ESA+South American clade; orange = the Hybridus clade; dark blue = the Dioecious/Pumilus clade; light green = the Galápagos clades.

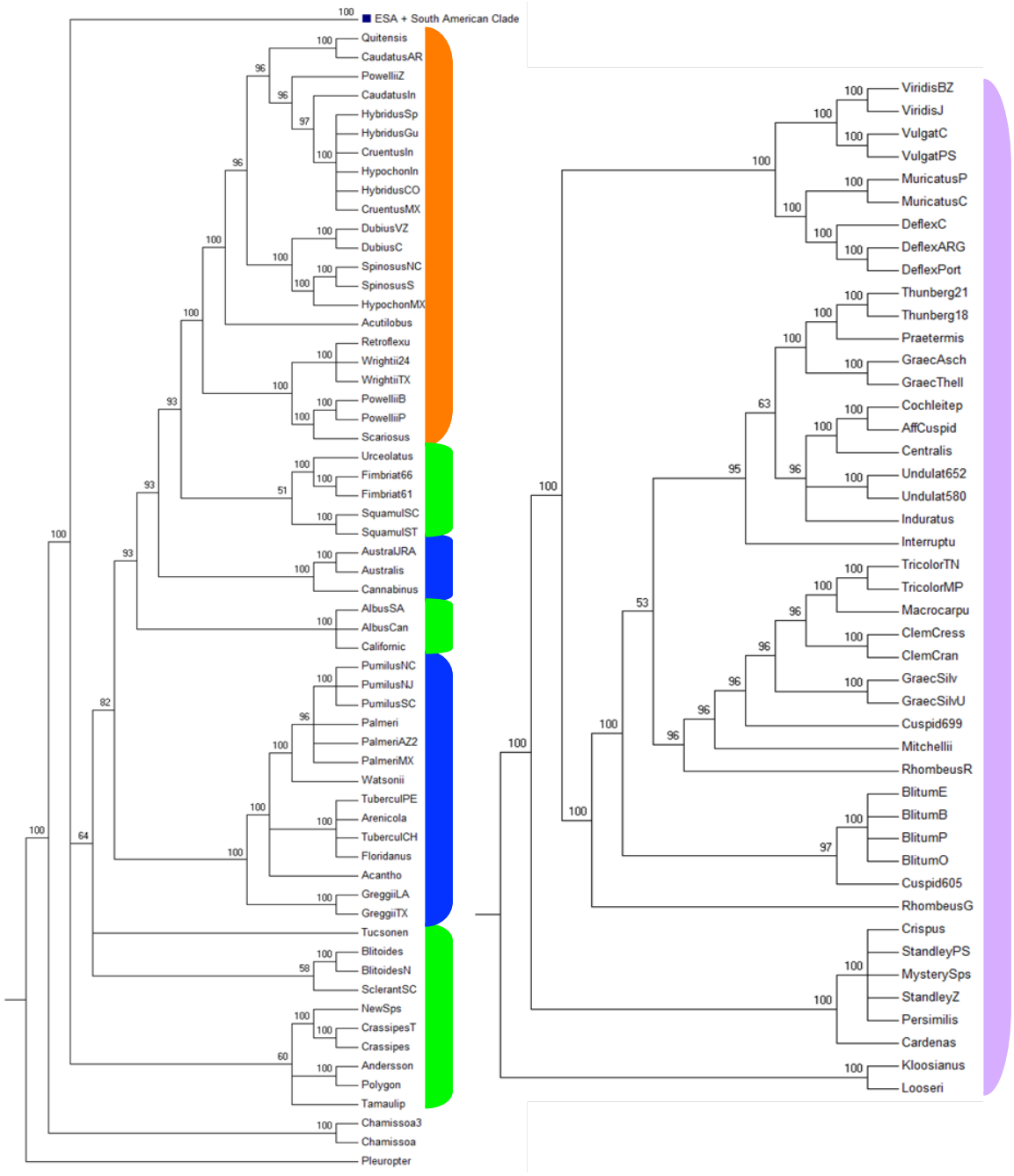

Figure 1.4. Maximum parsimony 50% majority-rule consensus tree of the most-parsimonious trees for the concatenated chloroplast dataset. The right-hand side shows the ESA + South American clade that has been collapsed in the left-hand tree. The major clades are shown with colored bars: Purple = the ESA+South American clade; orange = the Hybridus clade; dark blue = the Dioecious/Pumilus clades; light green = the Galápagos clades (plus extraneous species).

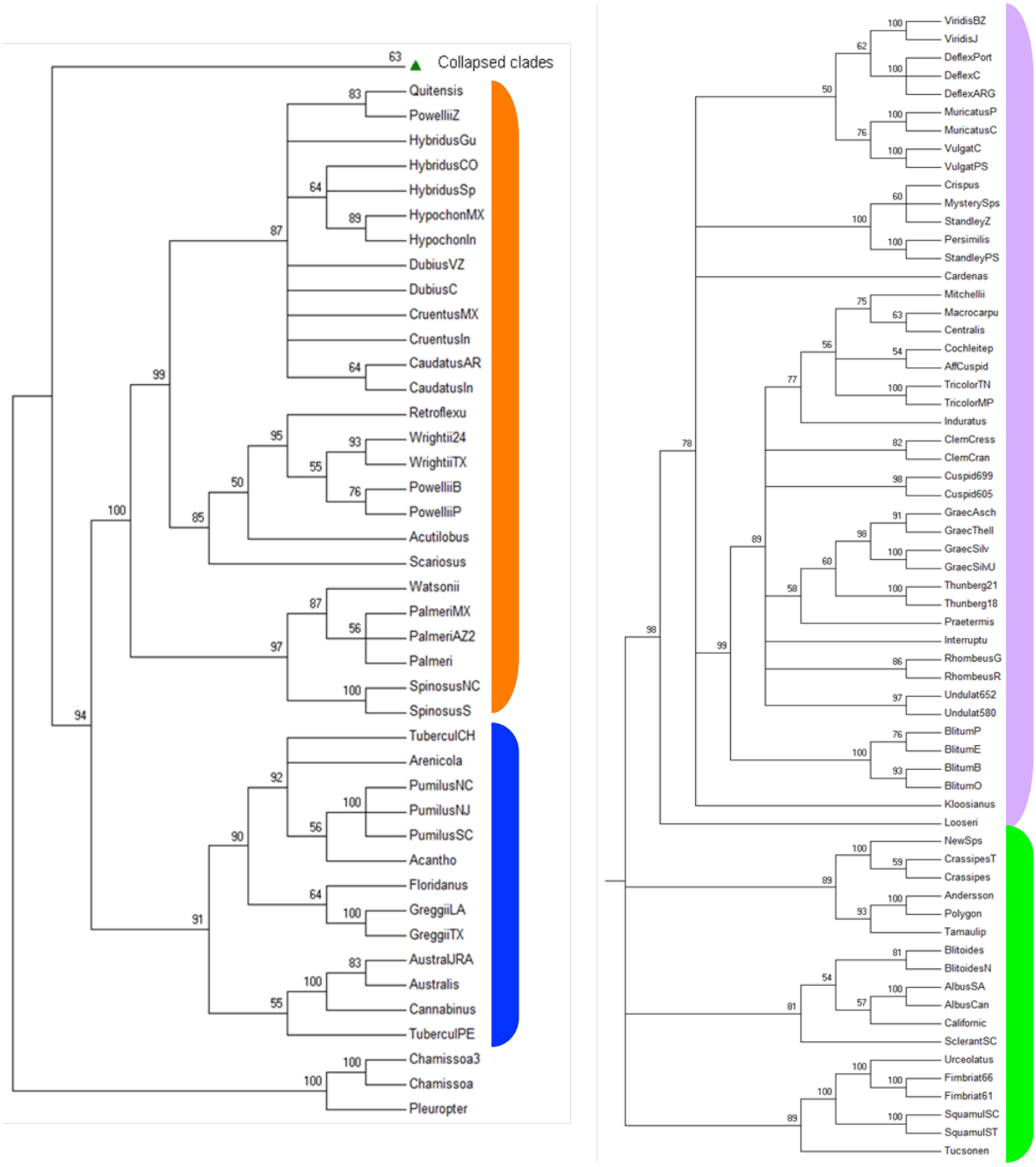

Figure 1.5. Maximum parsimony 50% majority-rule consensus tree from bootstrapping of the most-parsimonious trees for the concatenated nuclear dataset. Bootstrap support values are shown above the branches. The right-hand side shows the ESA + South American and Galápagos clades that have been collapsed in the left-hand tree. The major clades are shown with colored bars: Purple = the ESA+South American clade; orange = the Hybridus clade; dark blue = the Dioecious/Pumilus clade; light green = the Galápagos clades.

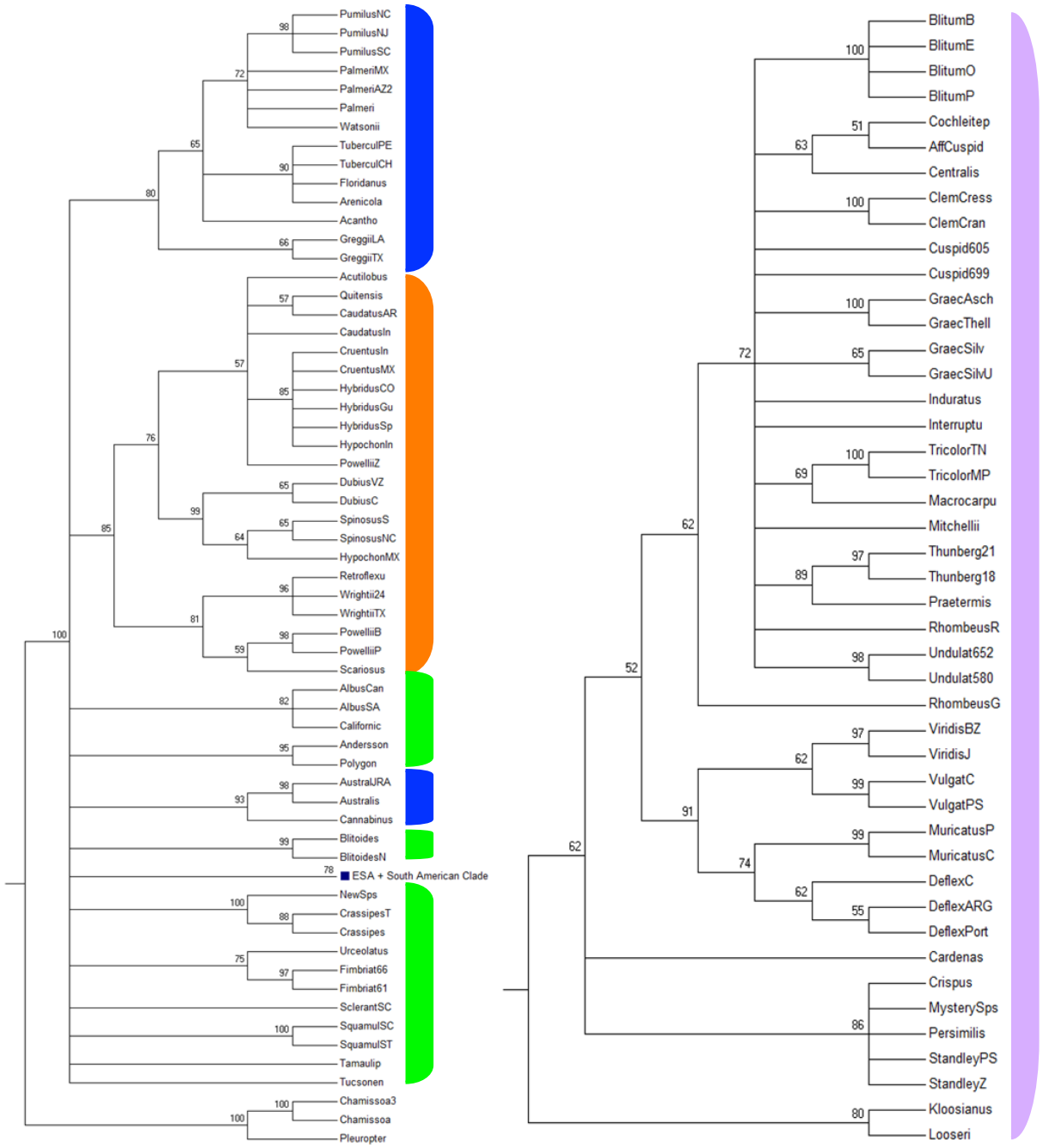

Figure 1.6. Maximum parsimony 50% majority-rule consensus tree from bootstrapping of the most-parsimonious trees for the concatenated chloroplast dataset. Bootstrap support values are shown above the branches. The right-hand side shows the ESA + South American clade that has been collapsed in the left-hand tree. The major clades are shown with colored bars: Purple = the ESA+South American clade; orange = the Hybridus clade; dark blue = the Dioecious/Pumilus clades; light green = the Galápagos clades (plus extraneous species).

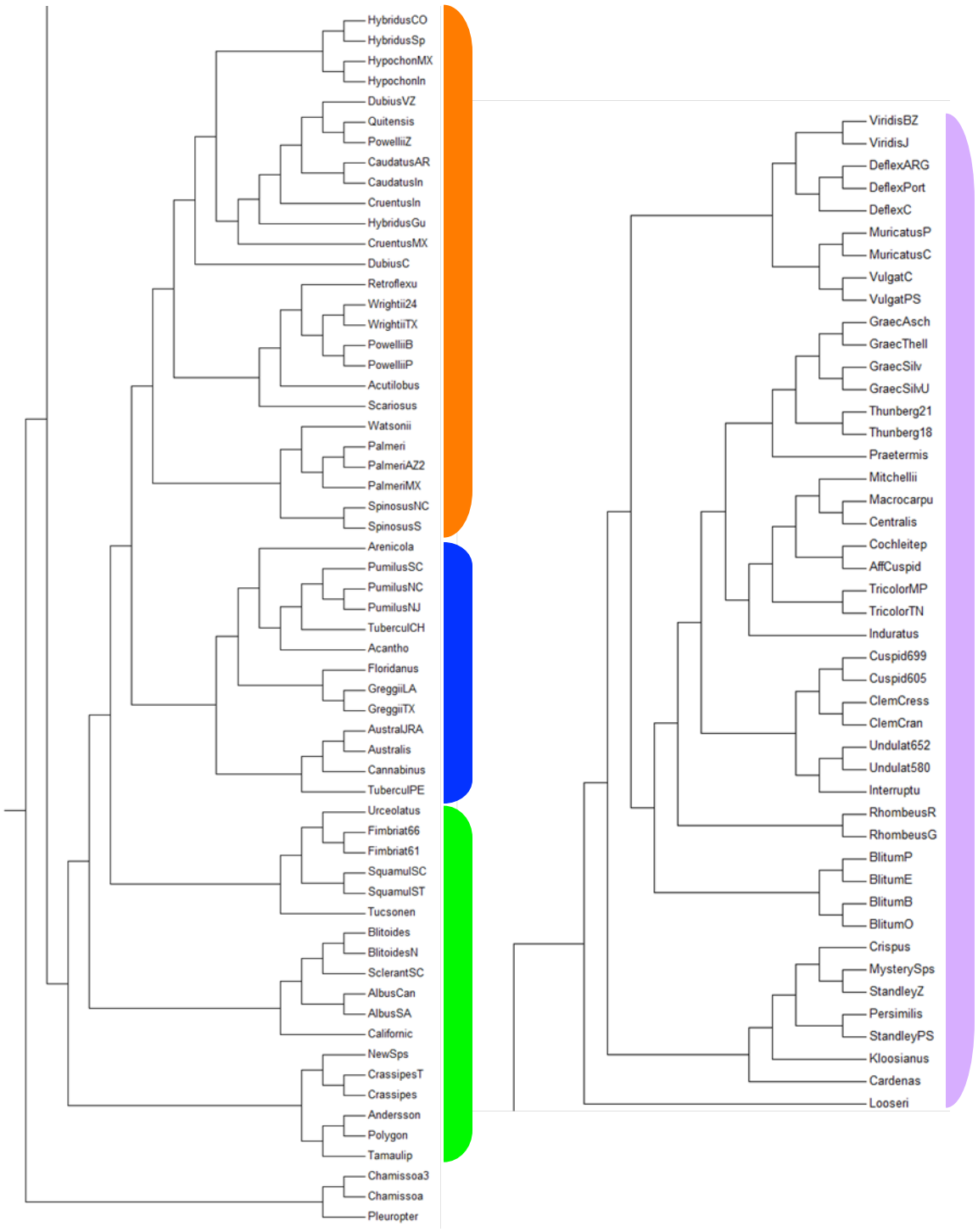

Figure 1.7. Maximum likelihood 50% majority-rule consensus tree of the highest-likelihood trees for the concatenated nuclear dataset. The base of the tree is on the left. The major clades are shown with colored bars: Purple = the ESA+South American clade; orange = the Hybridus clade; dark blue = the Dioecious/Pumilus clade; light green = the Galápagos clades.

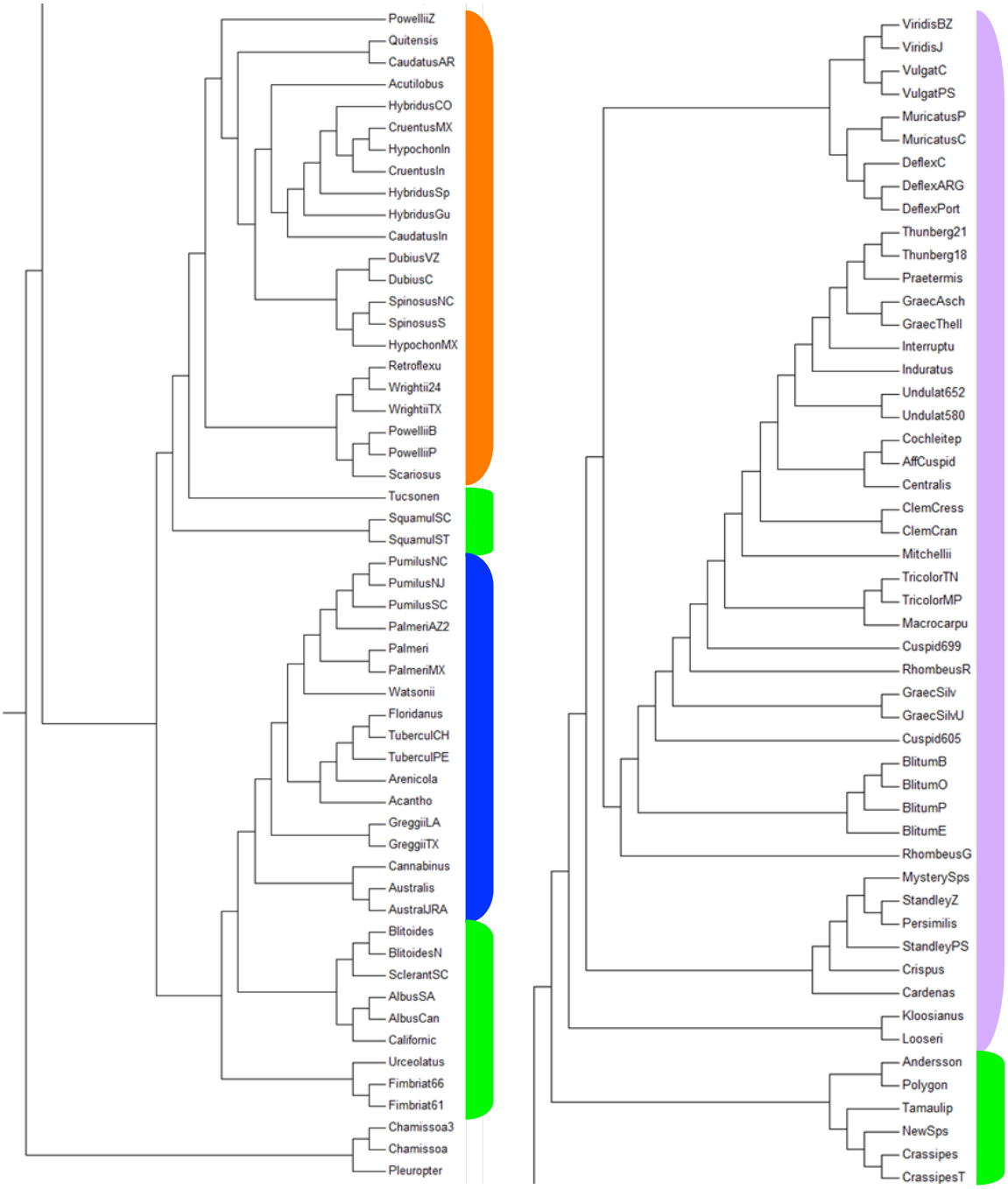

Figure 1.8. Maximum likelihood 50% majority-rule consensus tree of the highest-likelihood trees for the concatenated chloroplast dataset. The base of the tree is on the left. The major clades are shown with colored bars: Purple = the ESA+South American clade; orange = the Hybridus clade; dark blue = the Dioecious/Pumilus clade; light green = the Galápagos clades (plus extraneous species).

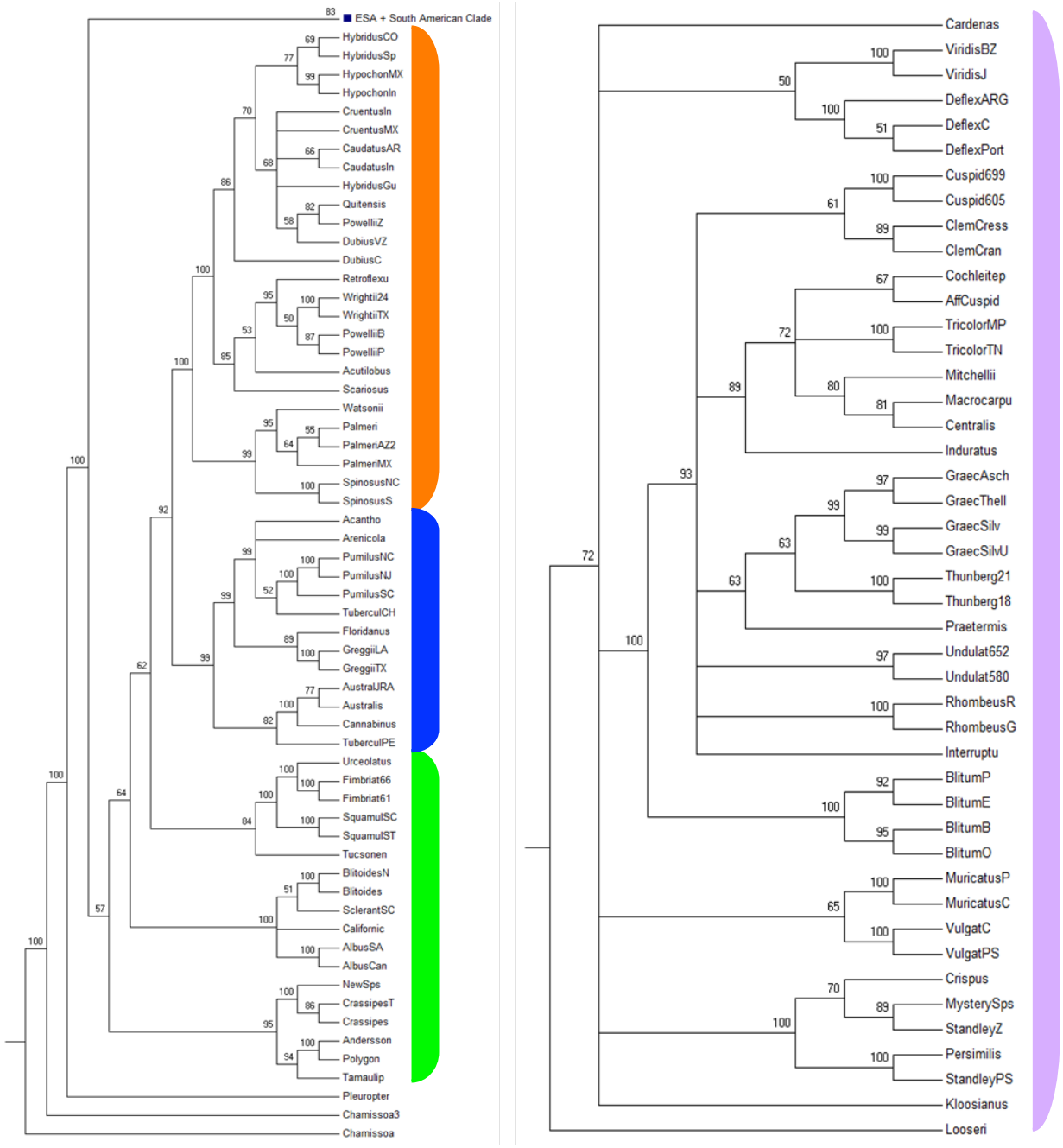

Figure 1.9. Maximum likelihood 50% majority-rule consensus tree from bootstrapping of the highest-likelihood trees for the concatenated nuclear dataset. Bootstrap support values are shown above the branches. The right-hand side shows the ESA + South American clade that has been collapsed in the left-hand tree. The major clades are shown with colored bars: Purple = the ESA+South American clade; orange = the Hybridus clade; dark blue = the Dioecious/Pumilus clade; light green = the Galápagos clades.

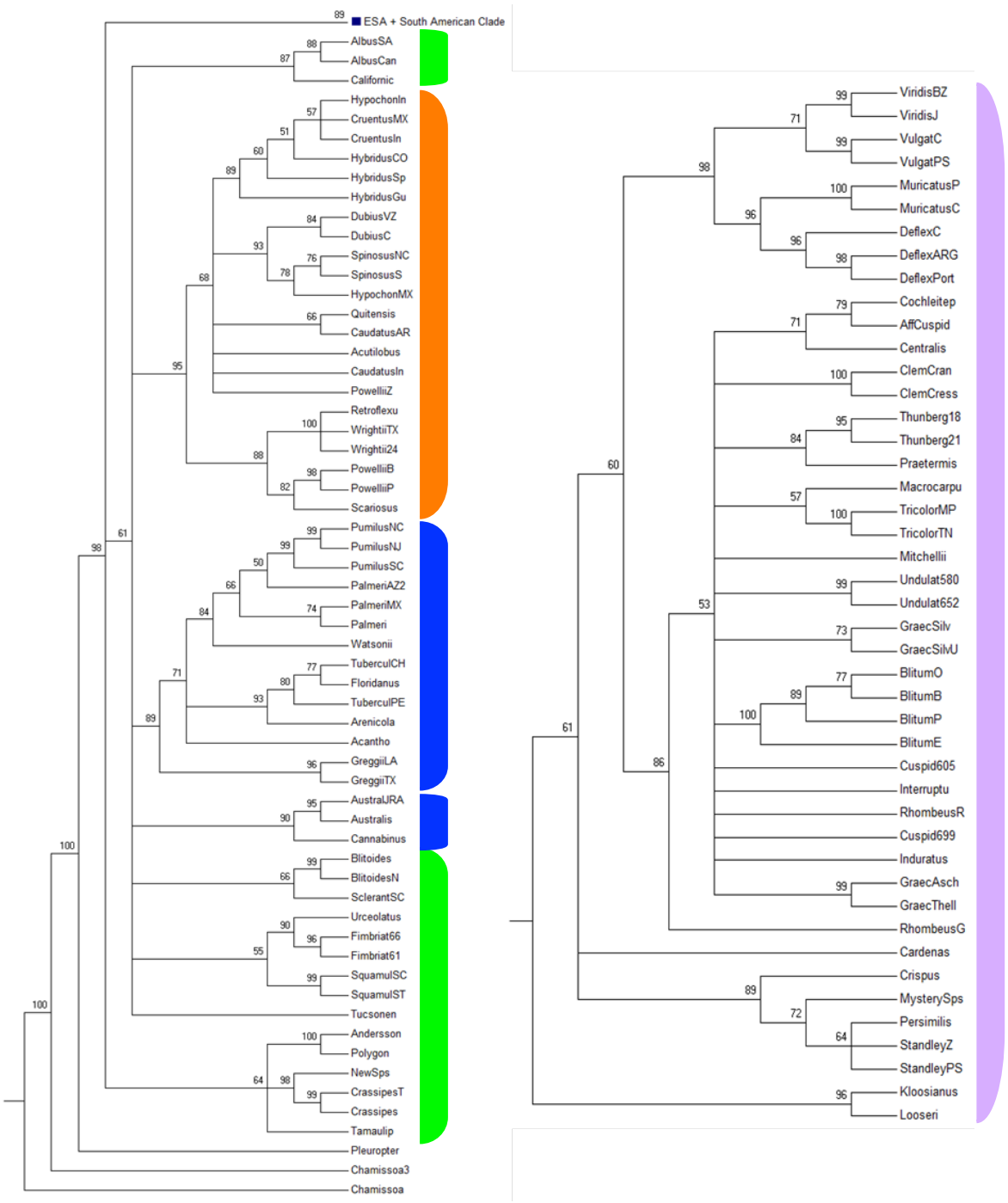

Figure 1.10. Maximum likelihood 50% majority-rule consensus tree from bootstrapping of the highest-likelihood trees for the concatenated chloroplast dataset. Bootstrap support values are shown above the branches. The right-hand side shows the ESA + South American clade that has been collapsed in the left-hand tree. The major clades are shown with colored bars: Purple = the ESA+South American clade; orange = the Hybridus clade; dark blue = the Dioecious/Pumilus clades; light green = the Galápagos clades (plus extraneous species).

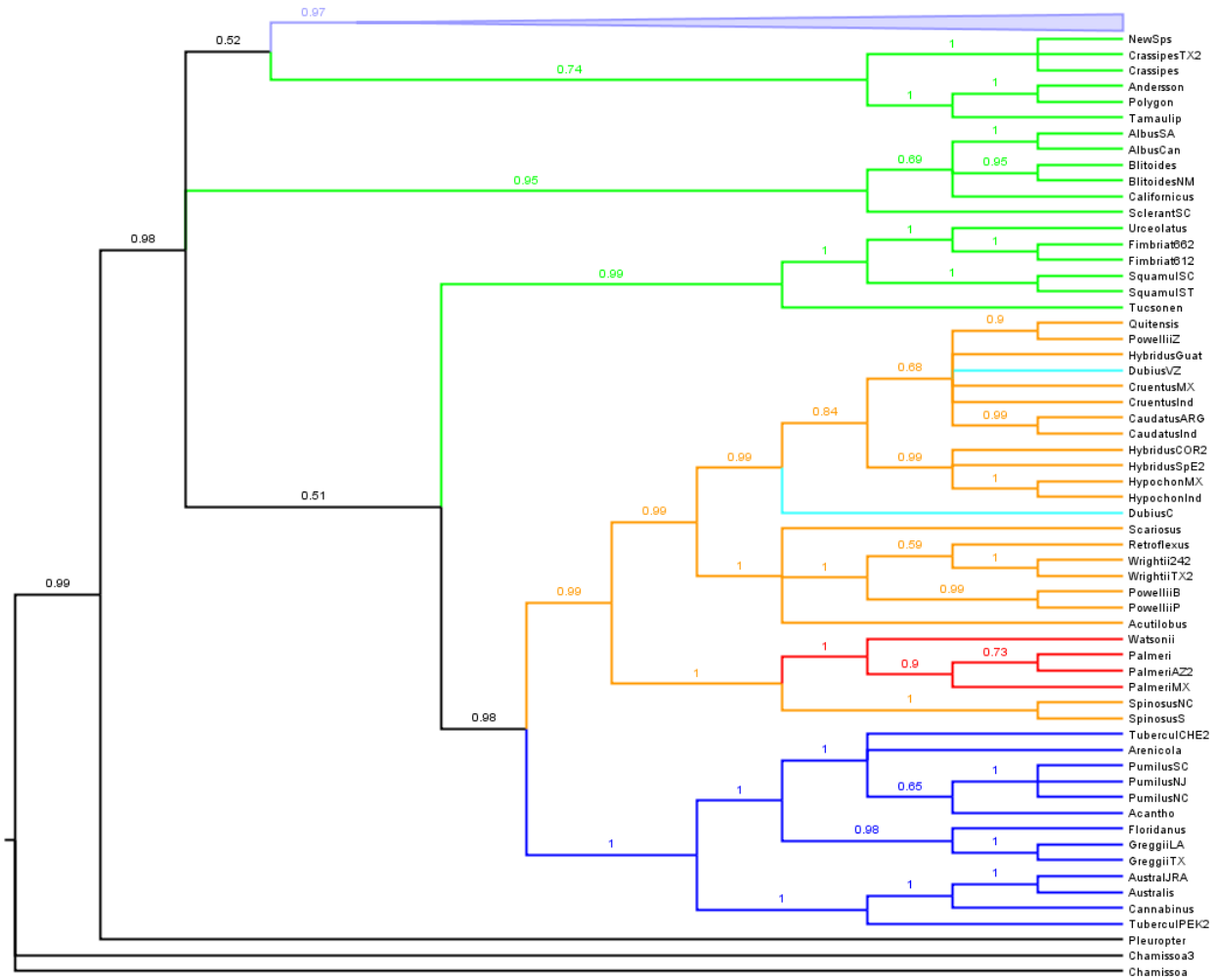

Figure 1.11a. Bayesian 50% majority-rule consensus tree with posterior probability values for the partitioned model for the concatenated nuclear dataset, showing the major clades and the areas of major disagreement with the concatenated chloroplast tree in different colors. Purple = the ESA+South American clade; orange = the Hybridus clade; dark blue = the Dioecious/Pumilus clade; light green = the Galápagos clades; red = *A. palmeri* and *A. watsonii*; light blue = *A. dubius*. Figure 1.11b shows the collapsed clade at the top of the tree.

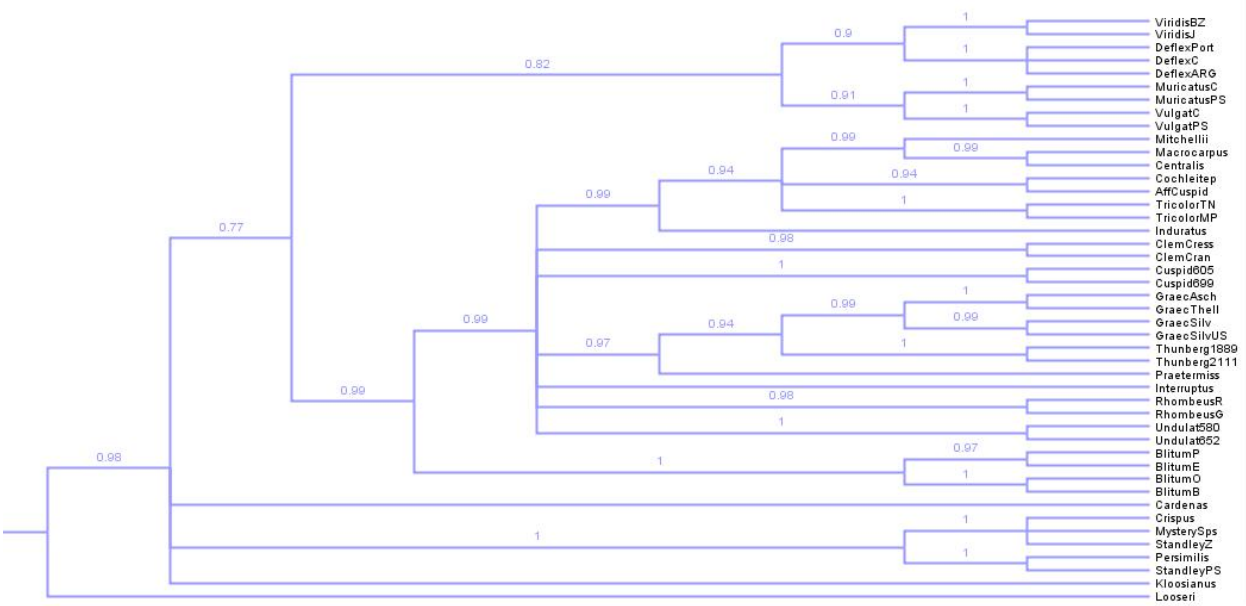

Figure 1.11b. The ESA+South American clade collapsed in Figure 1.11a.

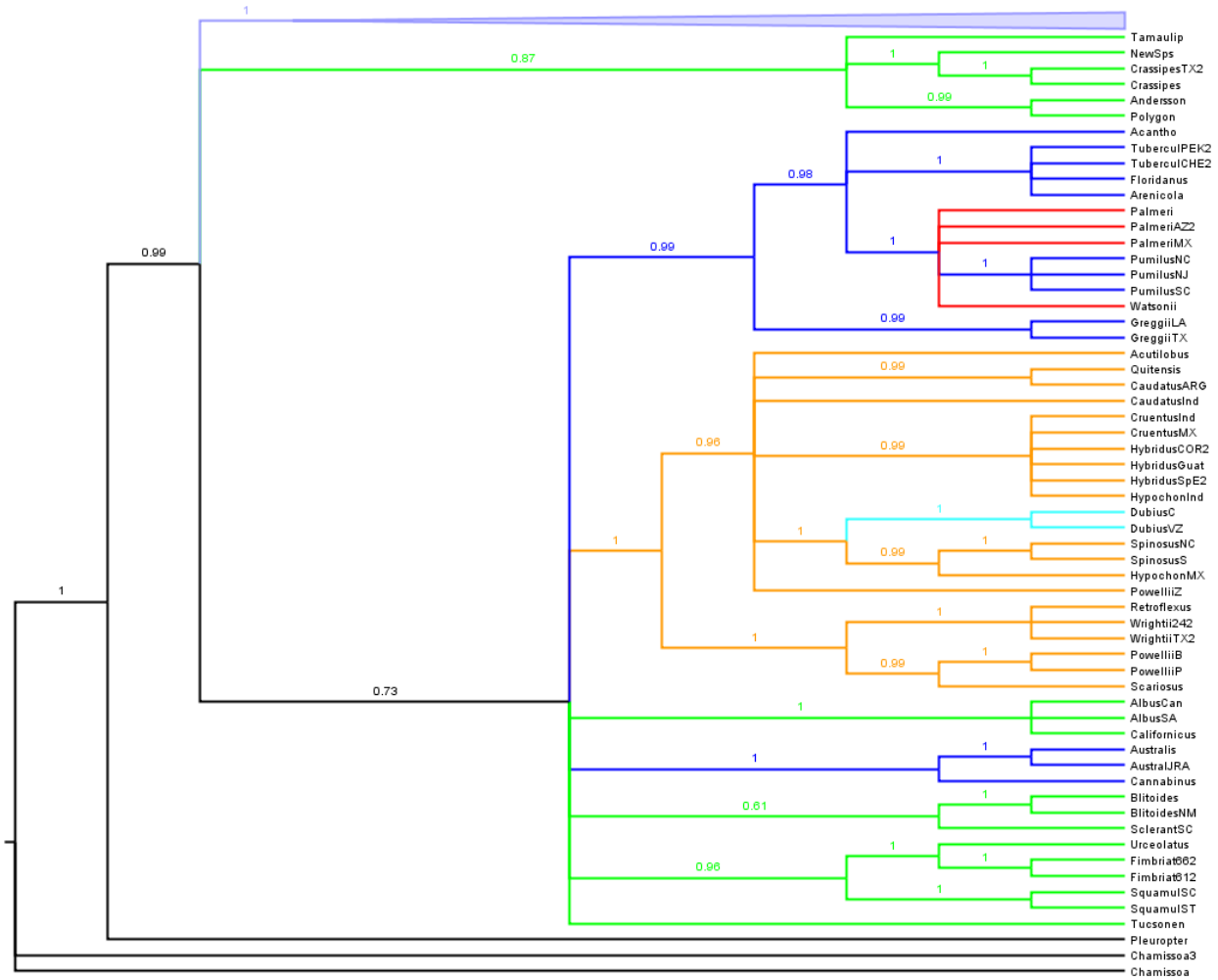

Figure 1.12a. Bayesian 50% majority-rule consensus tree with posterior probability values for the GTR+Γ model for the concatenated chloroplast dataset, showing the major clades and the areas of major disagreement with the concatenated chloroplast tree in different colors. Purple = the ESA+South American clade; orange = the Hybridus clade; dark blue = the Dioecious/Pumilus clade; light green = the Galápagos clades; red = *A. palmeri* and *A. watsonii*; light blue = *A. dubius*. Figure 1.12b shows the collapsed clade at the top of the tree.

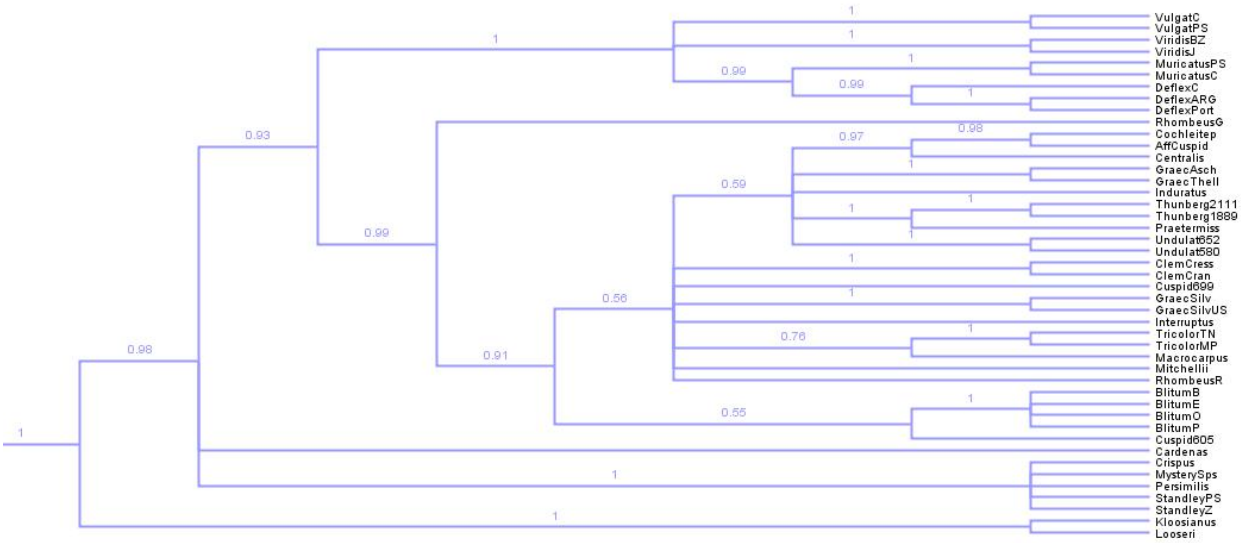

Figure 1.12b. The ESA+South American clade collapsed in Figure 1.12a.

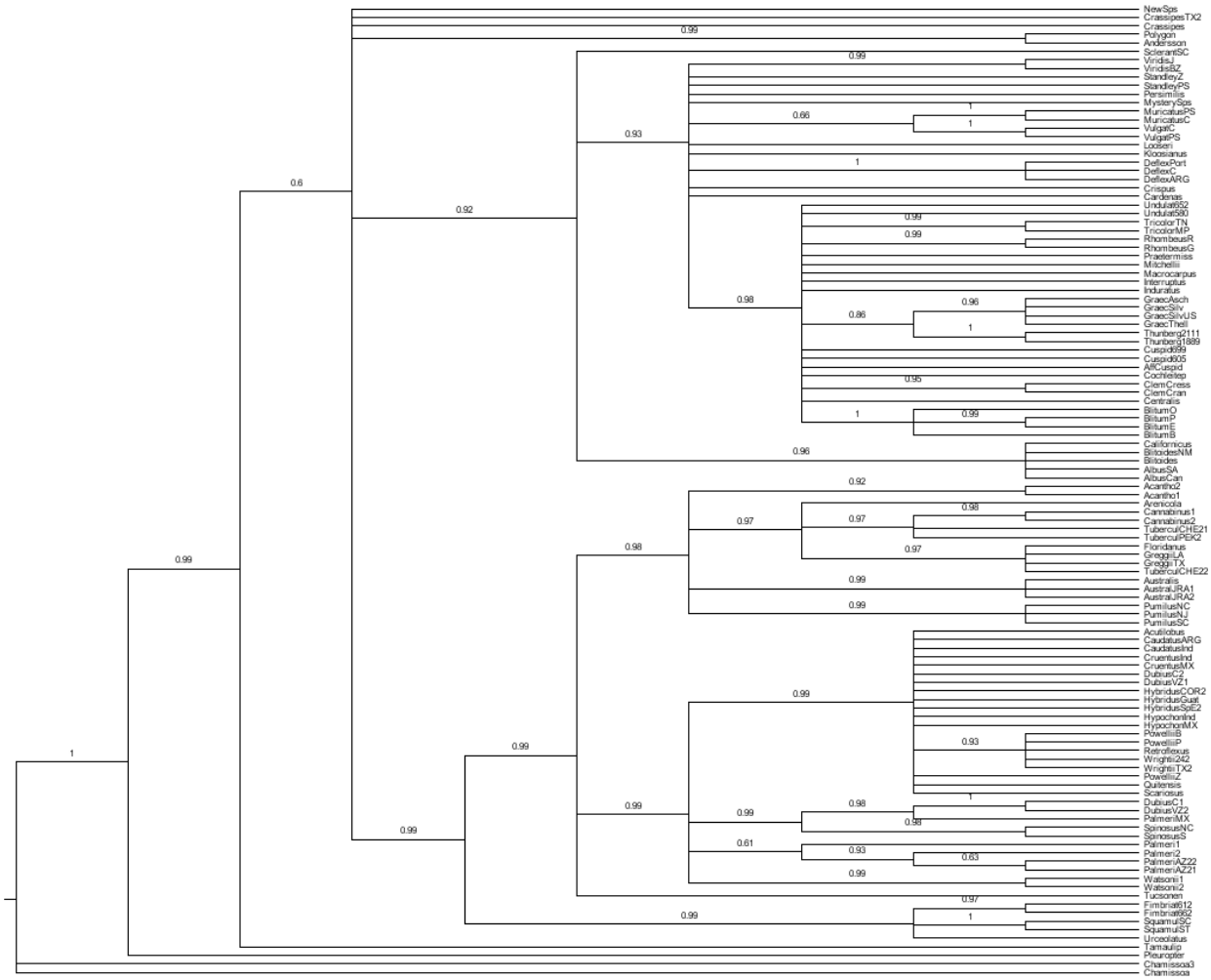

Figure 1.13a. Bayesian 50% majority-rule consensus tree with posterior probability values for the GTR+Γ model for the A36 gene.

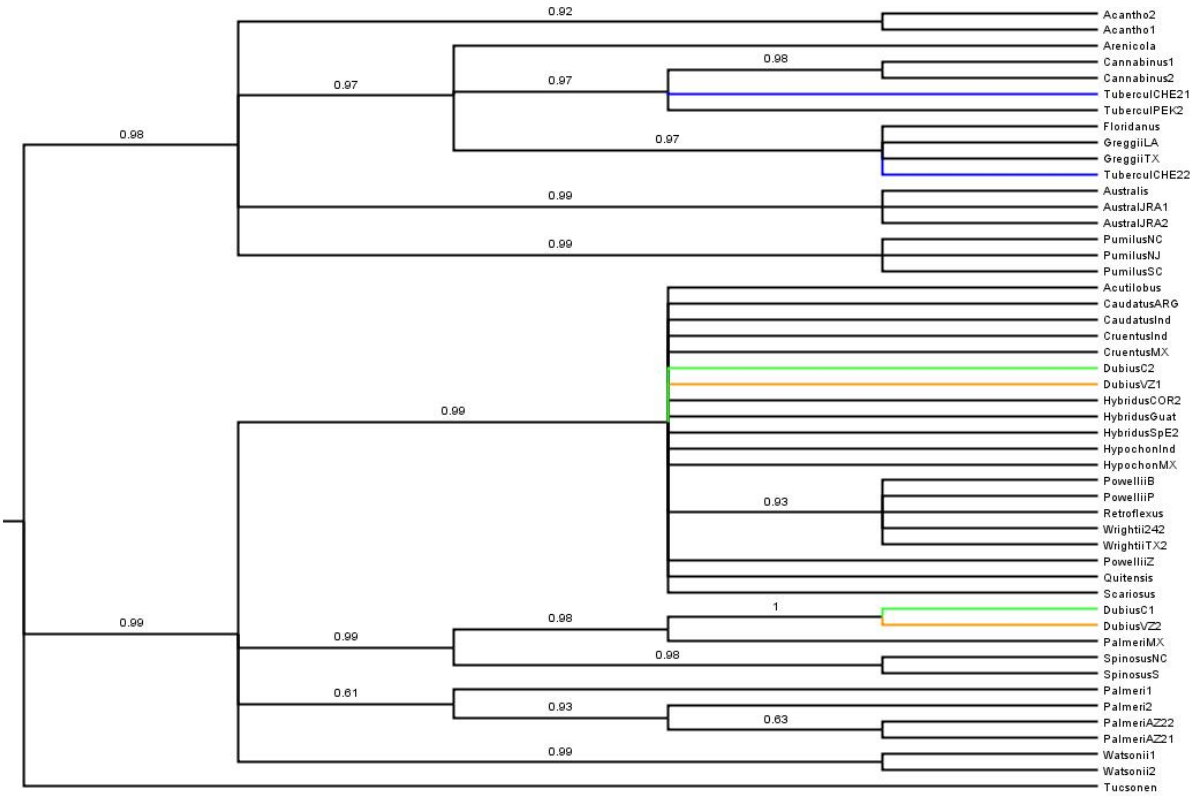

Figure 1.13b. A portion of the Bayesian 50% majority-rule consensus tree with posterior probability values for the GTR+Γ model for the A36 gene, showing incomplete lineage sorting. Alleles of the same accession that are highly supported as non-monophyletic are in the same color.

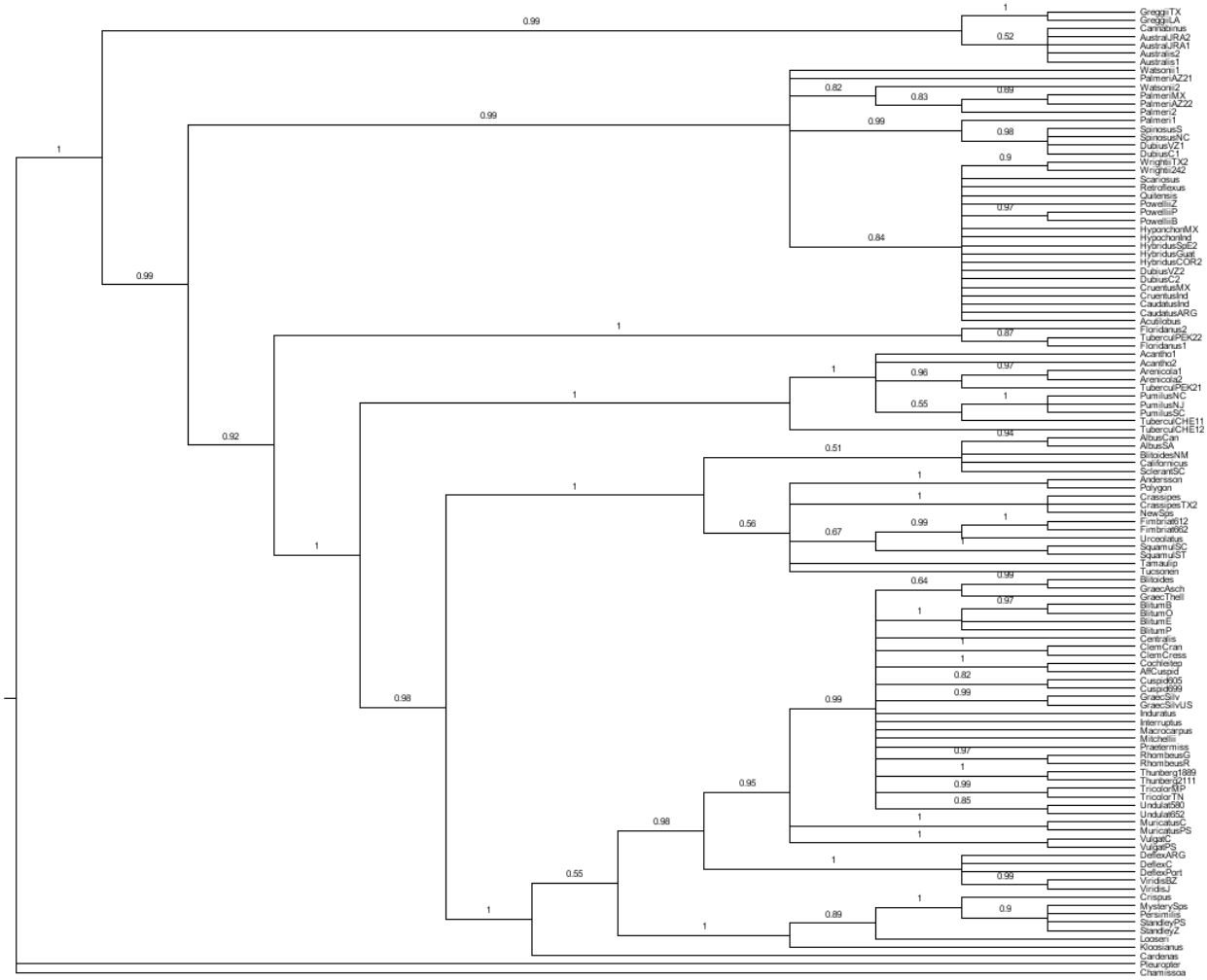

Figure 1.14a. Bayesian 50% majority-rule consensus tree with posterior probability values for the GTR+Γ model for the G3PDH gene. *Amaranthus* is constrained to be monophyletic.

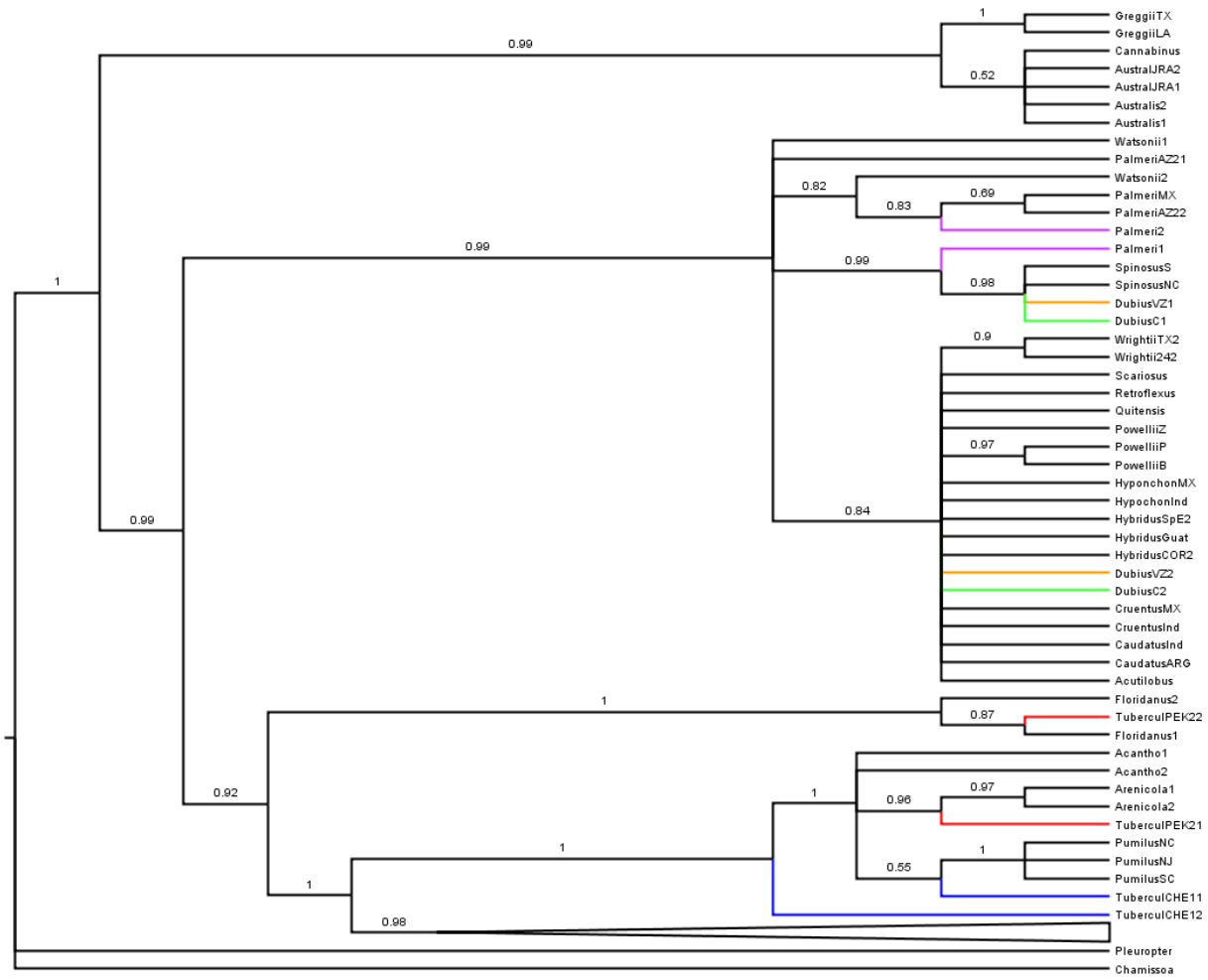

Figure 1.14b. A portion of the Bayesian 50% majority-rule consensus tree with posterior probability values for the GTR+Γ model for the G3PDH gene, showing incomplete lineage sorting. Alleles of the same accession that are highly supported as non-monophyletic are in the same color.

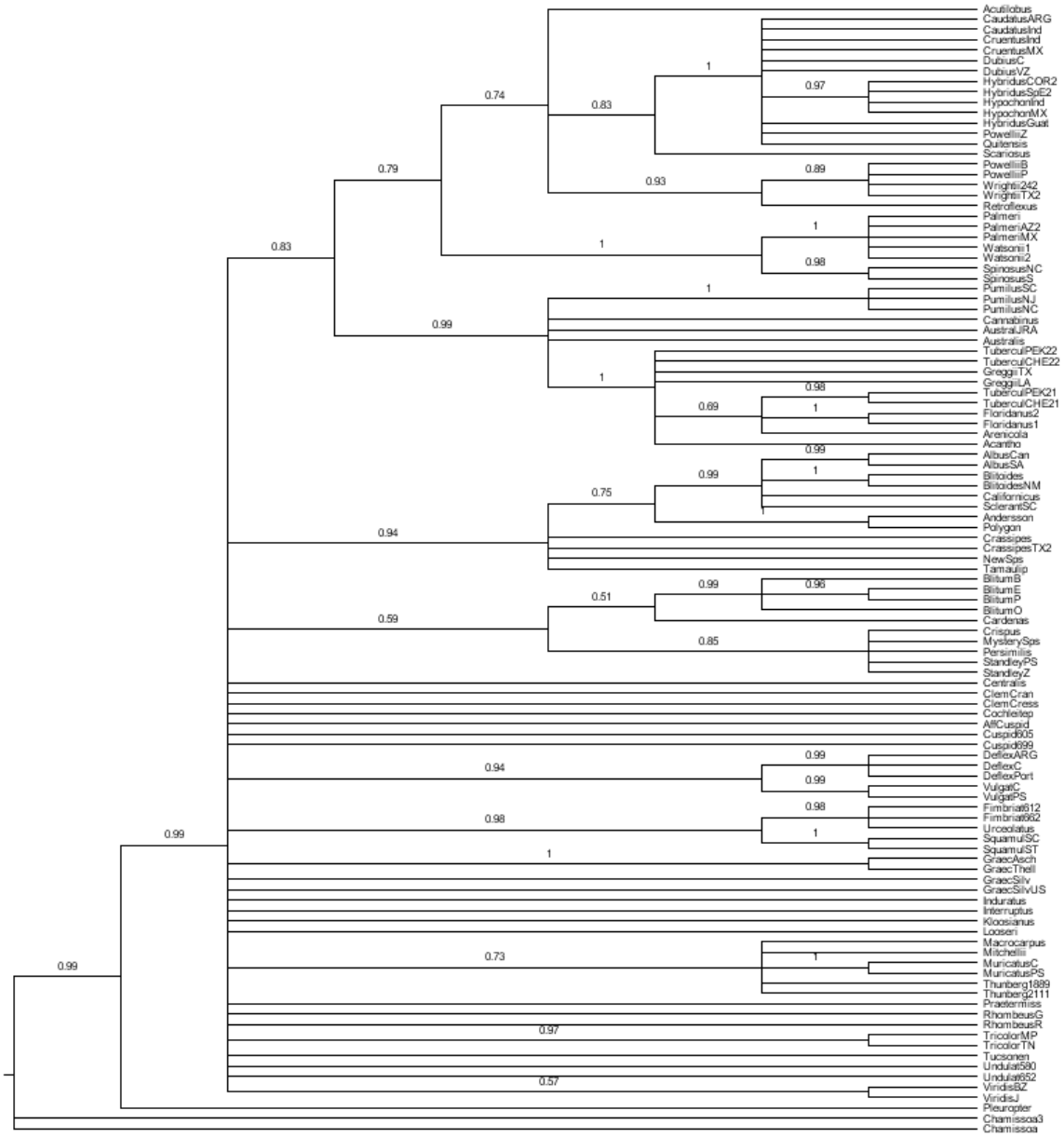

Figure 1.15a. Bayesian 50% majority-rule consensus tree with posterior probability values for the GTR+I+Γ model for the ITS gene.

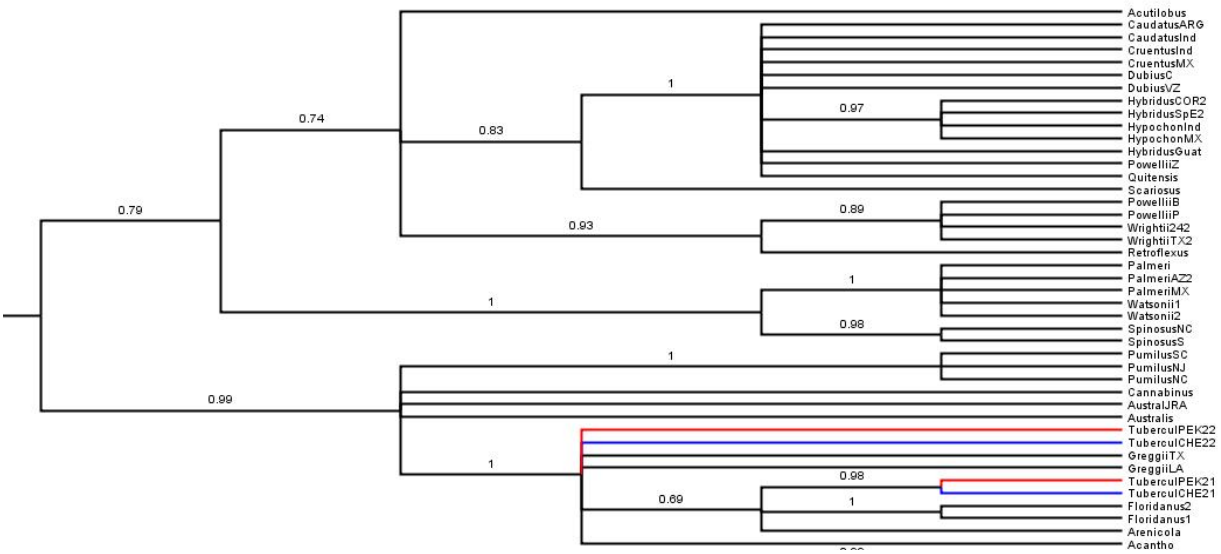

Figure 1.15b. A portion of the Bayesian 50% majority-rule consensus tree with posterior probability values for the GTR+I+Γ model for the ITS gene, showing incomplete lineage sorting. Alleles of the same accession that are highly supported as non-monophyletic are in the same color.

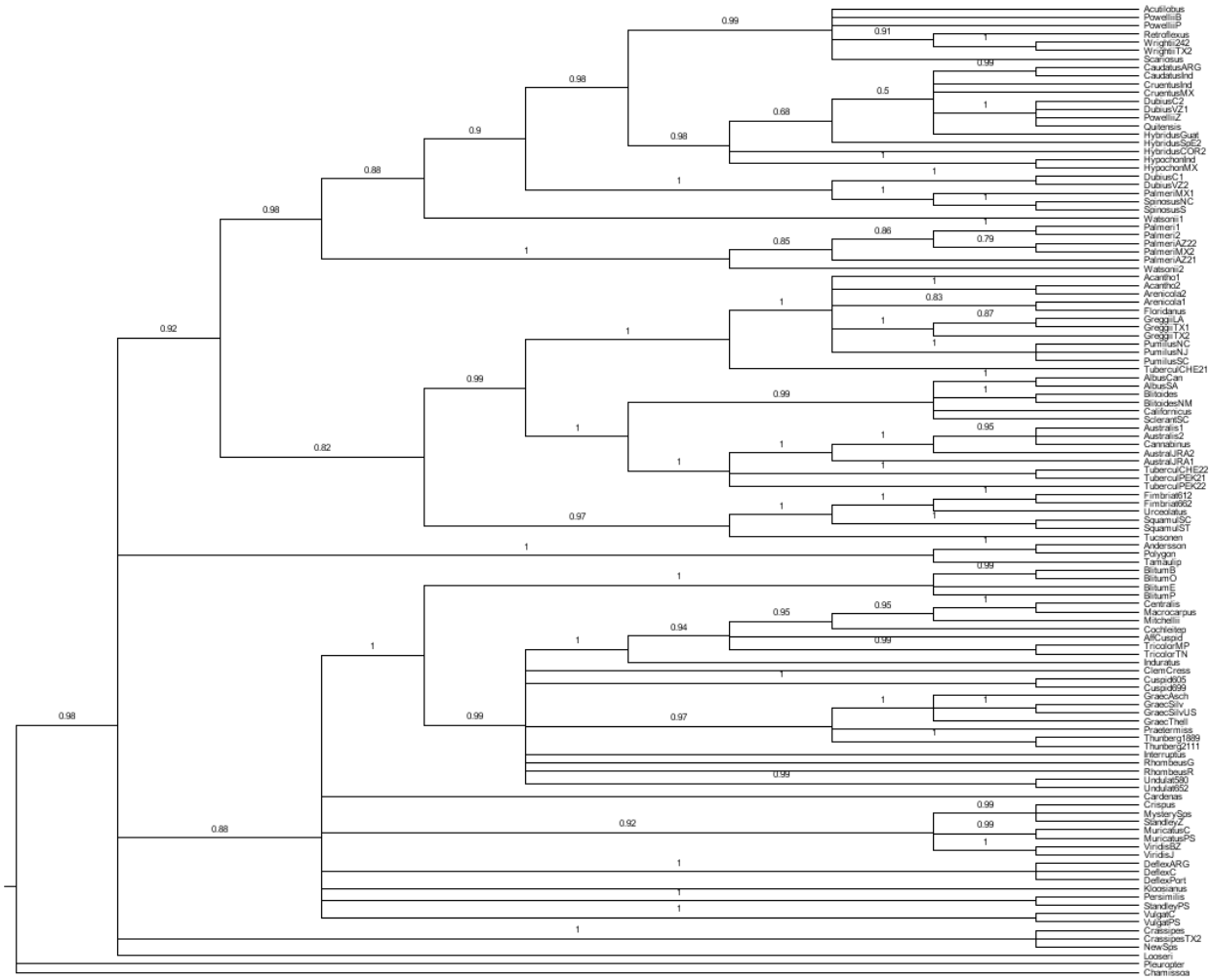

Figure 1.16a. Bayesian 50% majority-rule consensus tree with posterior probability values for the HKY+Γ model for the Waxy gene.

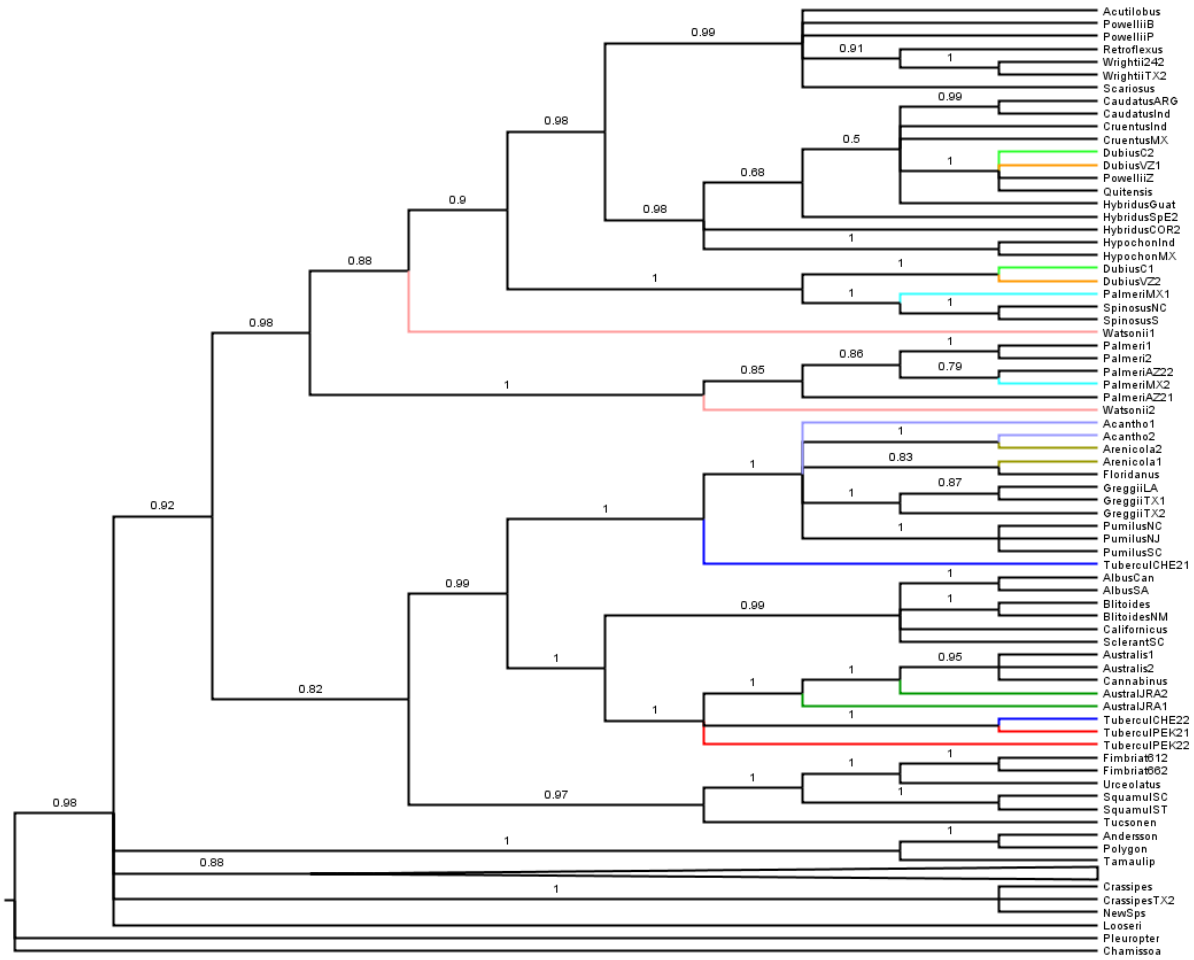

Figure 1.16b. A portion of the Bayesian 50% majority-rule consensus tree with posterior probability values for the HKY+Γ model for the Waxy gene, showing incomplete lineage sorting. Alleles of the same accession that are highly supported as non-monophyletic are in the same color.

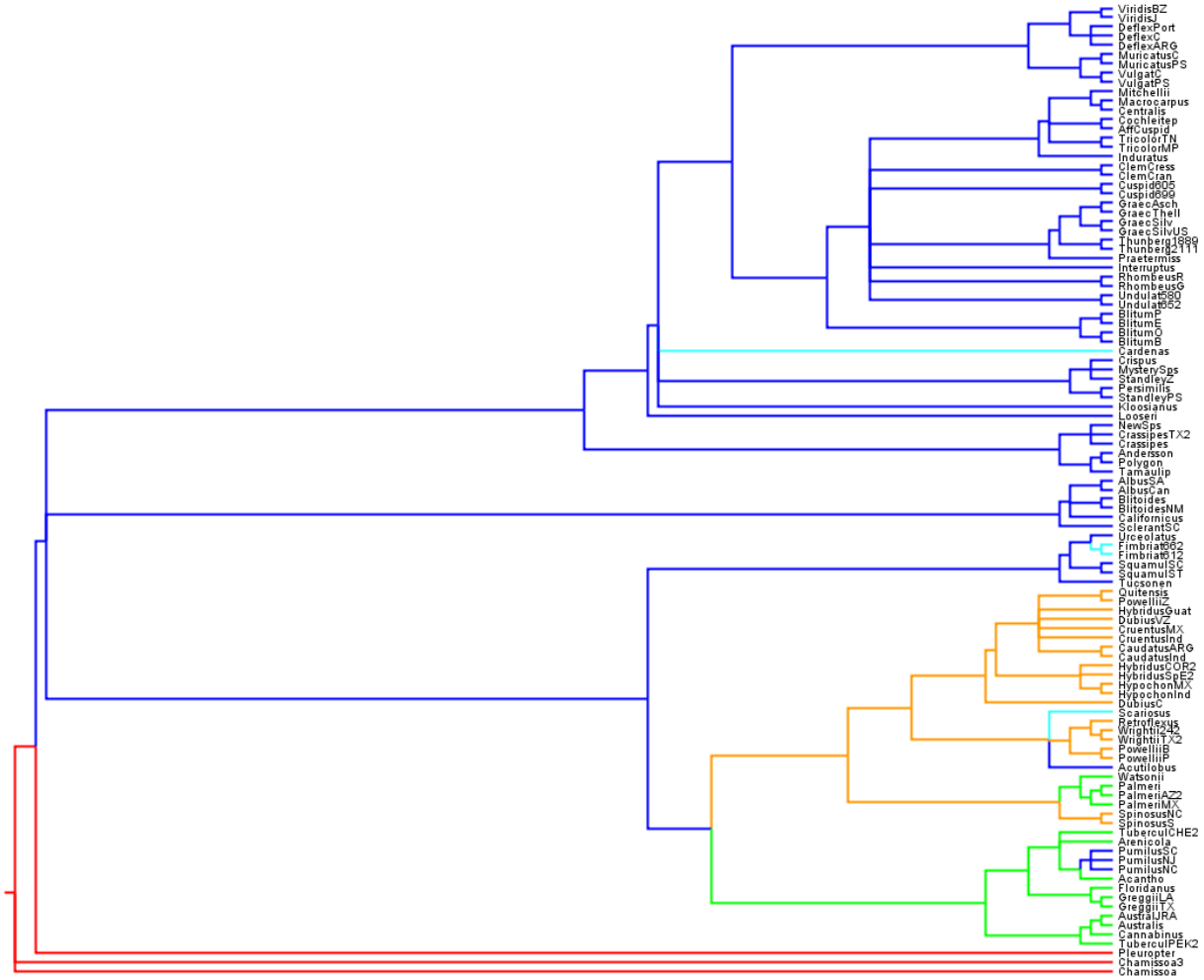

Figure 1.17. Bayesian 50% majority-rule consensus tree for the partitioned model for the concatenated nuclear dataset, showing the taxonomic subgenera of each species in different colors. The subgenus information is taken from Table 1.1. Dark blue = subgenus Albersia; orange = subgenus Amaranthus; light green = subgenus Acnida; light blue = disagreement between the two literature sources; red = outgroups.

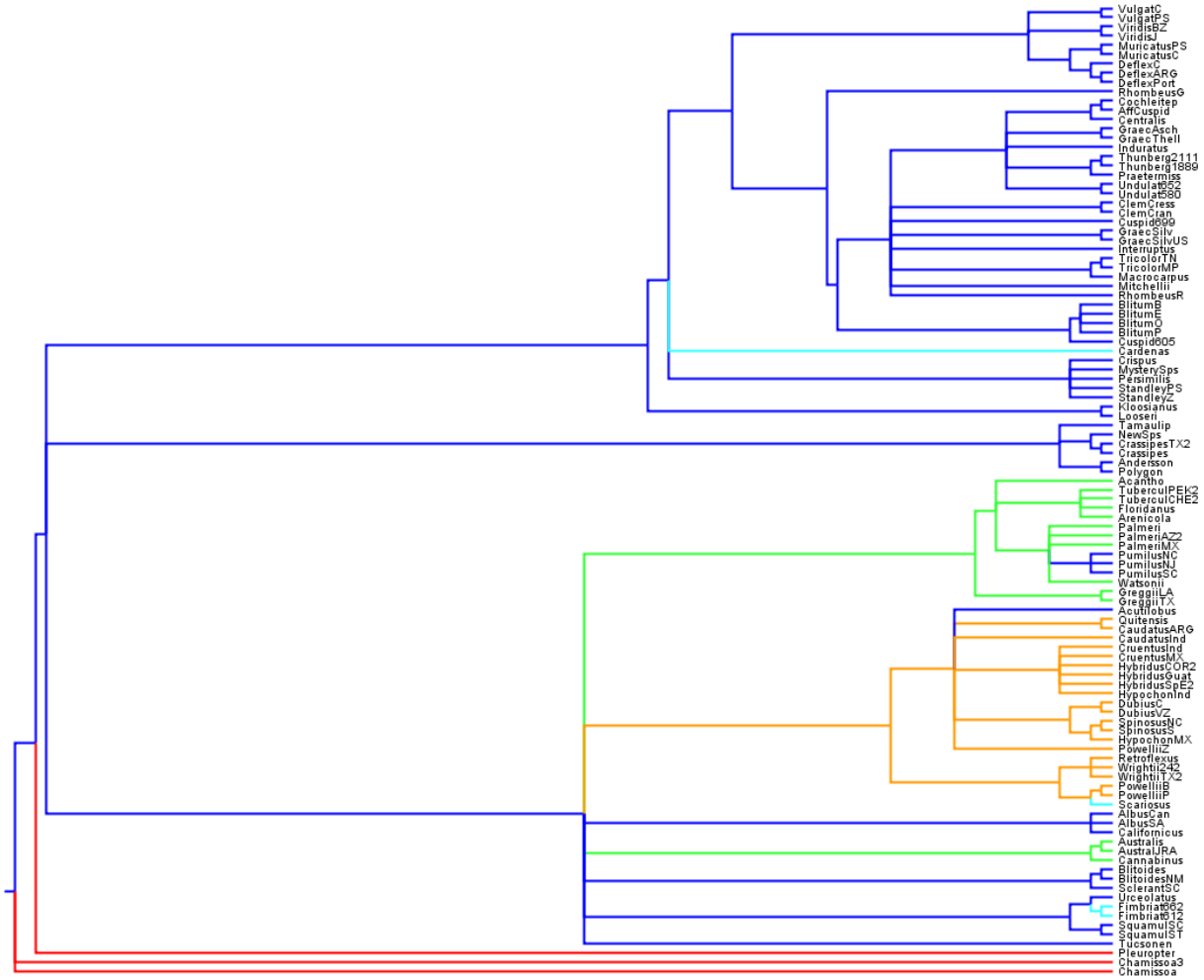

Figure 1.18. Bayesian 50% majority-rule consensus tree for the GTR+Γ model for the concatenated chloroplast dataset, showing the taxonomic subgenera of each species in different colors. The subgenus information is taken from Table 1.1. Dark blue = subgenus Albersia; orange = subgenus Amaranthus; light green = subgenus Acnida; light blue = disagreement between the two literature sources; red = outgroups.

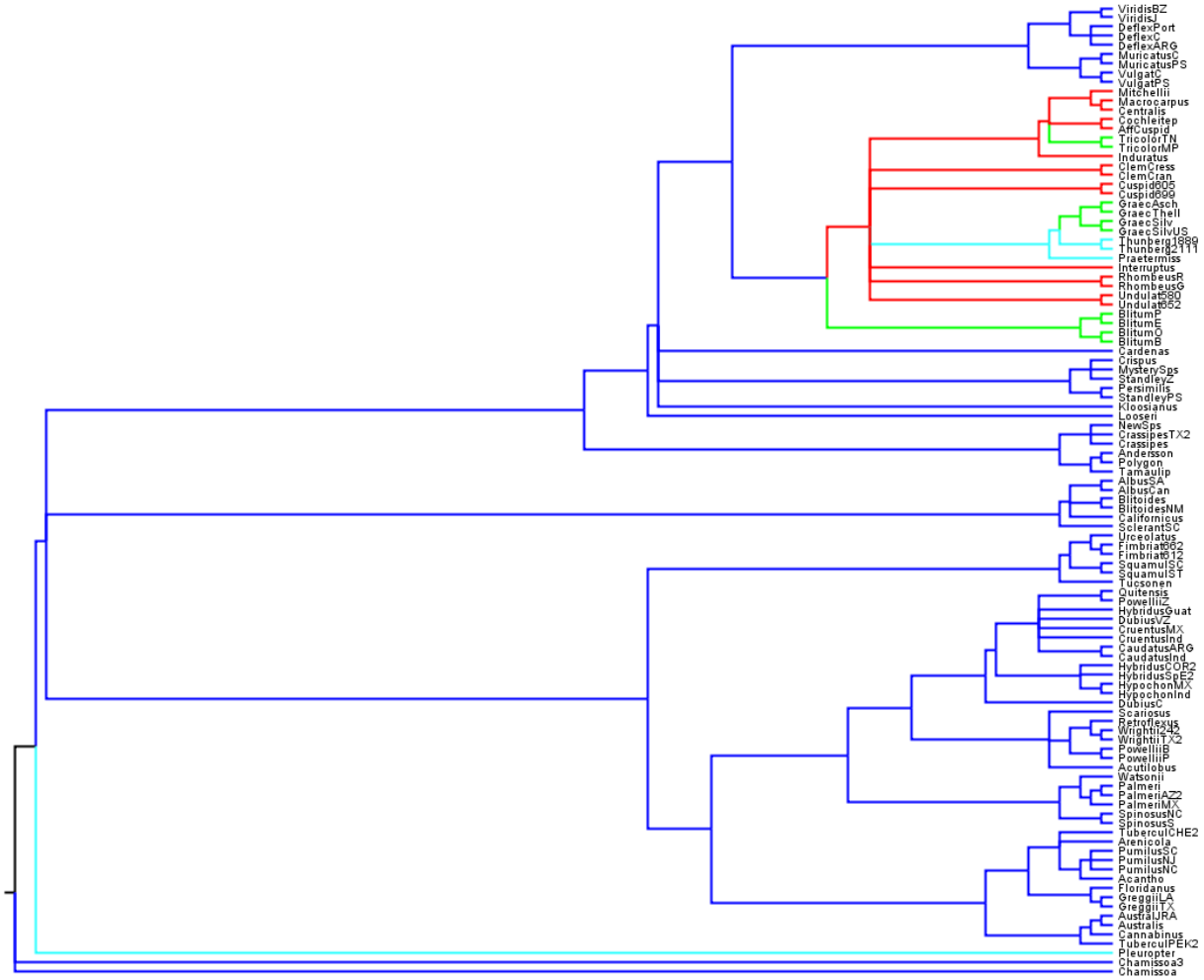

Figure 1.19. Bayesian 50% majority-rule consensus tree for the partitioned model for the concatenated nuclear dataset, showing the geographical origins of each species in different colors. Dark blue = the Americas; light blue = Africa; light green = Eurasia; red = Australia.

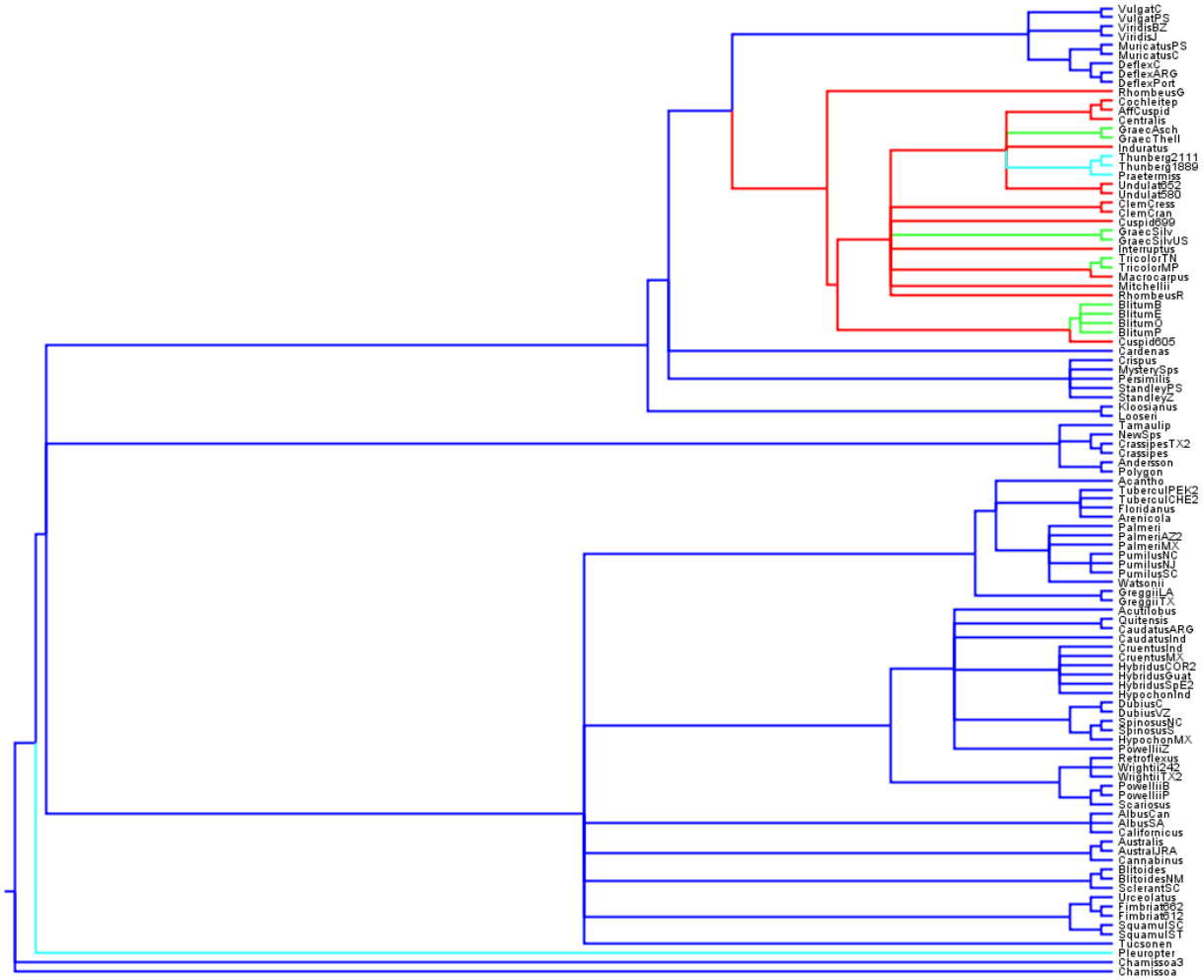

Figure 1.20. Bayesian 50% majority-rule consensus tree for the GTR+Γ model for the concatenated chloroplast dataset, showing the geographical origins of each species in different colors. Dark blue = the Americas; light blue = Africa; light green = Eurasia; red = Australia.

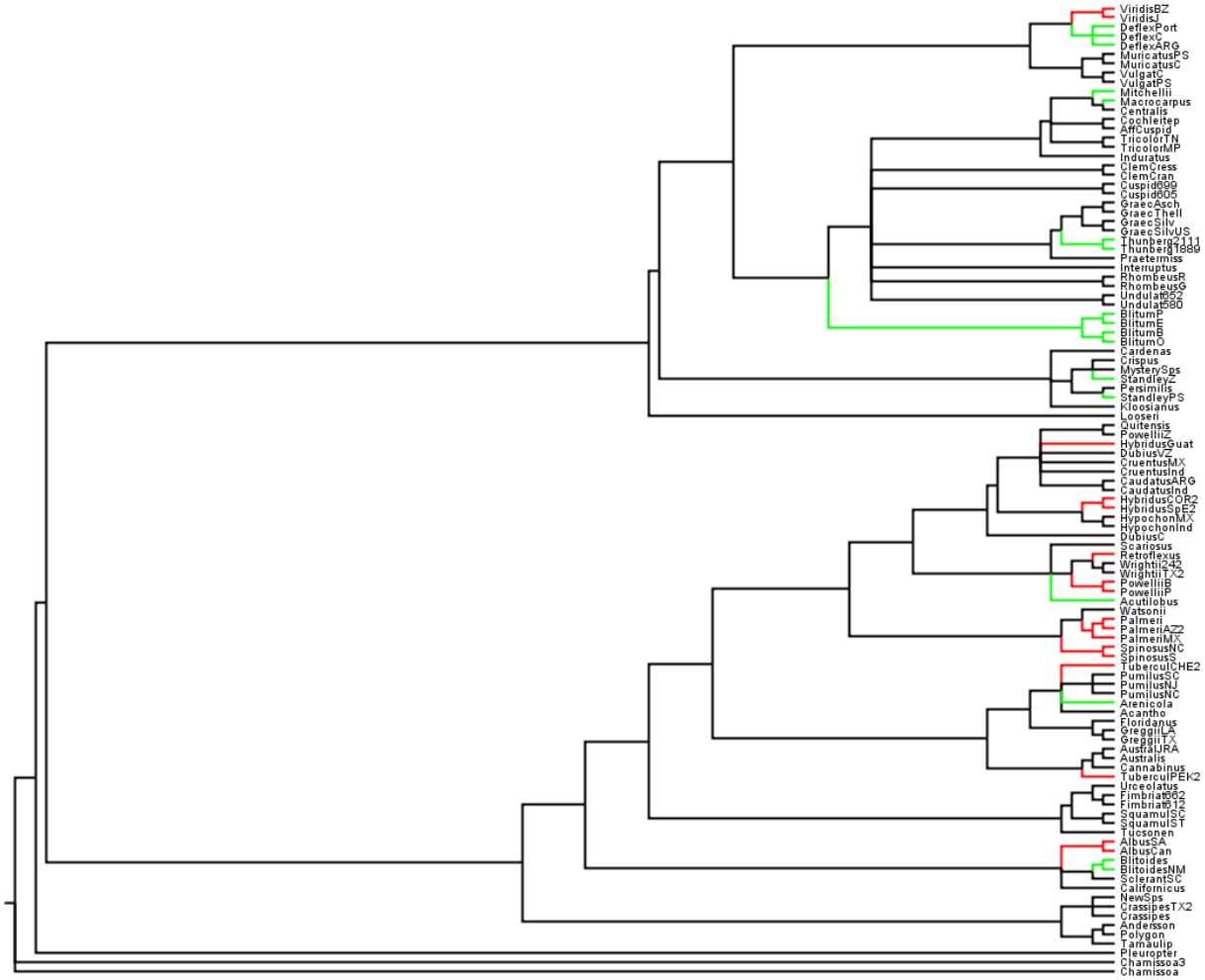

Figure 1.21. Bayesian 50% majority-rule consensus tree with posterior probability values for the partitioned model for the concatenated nuclear dataset, with weeds shown in color. Red = problematic weeds (Rank 3 in Table 1.5), light green = less problematic weeds.

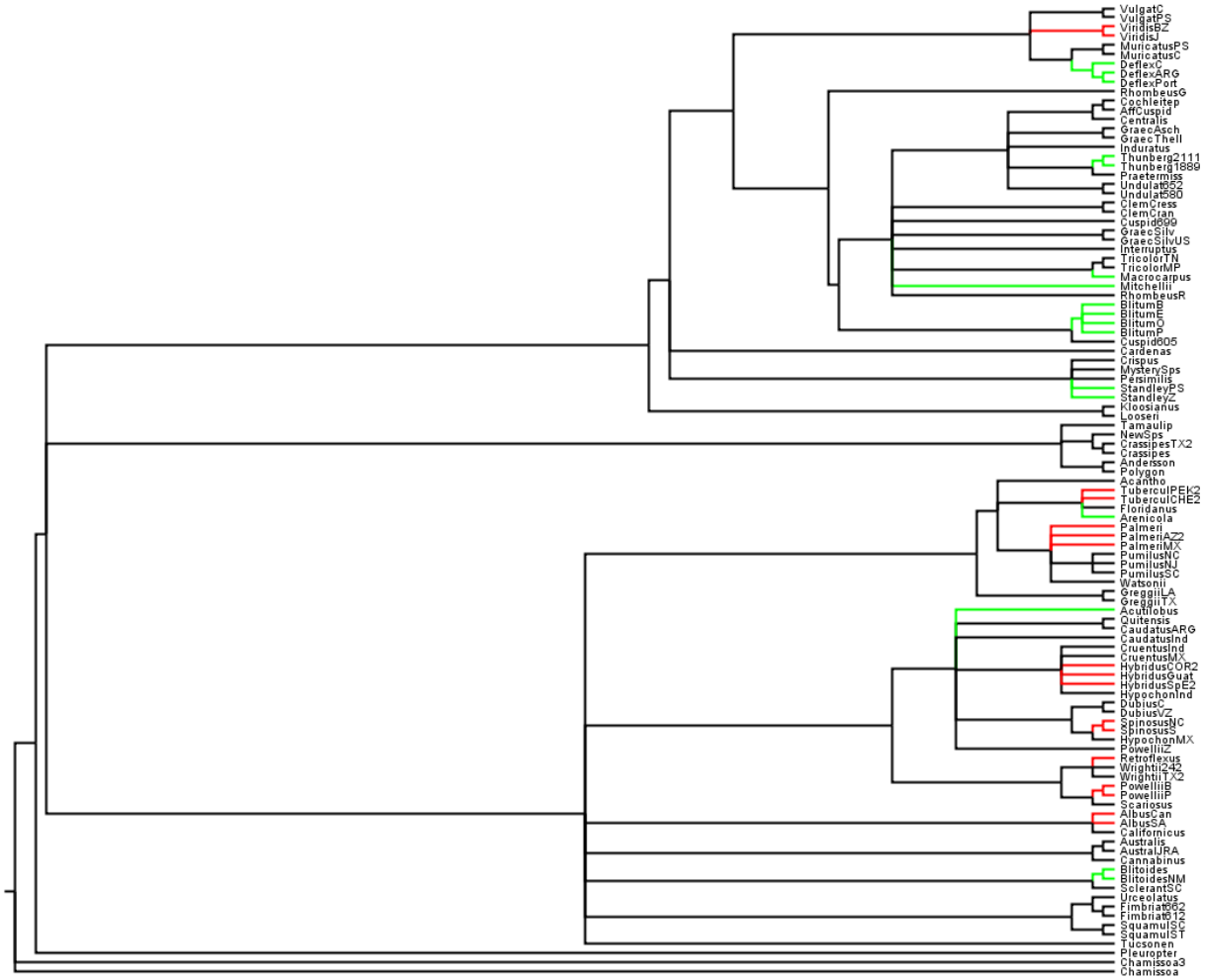

Figure 1.22. Bayesian 50% majority-rule consensus tree with posterior probability values for the GTR+Γ model for the concatenated chloroplast dataset, with weeds shown in color. Red = problematic weed species (Rank 3 in Table 1.5), light green = less problematic weeds (Rank 2 in Table 1.5).

CHAPTER 2

Population Genetics and Origin of the Native Midwestern Agricultural Weed, Waterhemp (*Amaranthus tuberculatus*)

INTRODUCTION

How does a plant species become invasive in agricultural ecosystems? Agricultural weeds are often presumed to have evolved along with plant domestication and the beginnings of agriculture (De Wet and Harlan, 1975). This ancient origin limits reconstruction of the evolutionary events that created them. However, agricultural weeds may also evolve on a more contemporary time frame. The 20th century saw enormous changes in agricultural practices in the U.S., including the introduction of herbicides and the widespread adoption of conservation tillage (Owen, 2008), and these changes may have allowed species which were formerly confined to natural habitats to find a new niche in an agricultural environment (Hilgenfeld et al., 2004). Unlike weeds whose origins date to the beginnings of agriculture, recently arisen weeds may retain a clear genetic signature of the events that led to their agricultural invasion.

Three main hypotheses about the origin of agricultural weeds are prevalent in the literature (reviewed in Vigueira et al., 2013, following De Wet and Harlan, 1975). Weed species that are related to domesticated species may arise either through "de-domestication" (domesticated species becoming feral), or by hybridization between related domesticated and wild species. Support for these hypotheses has been found in many systems, including beets, rye, rice, and sunflowers (Burger et al. 2006; Londo and Schaal, 2007; Olsen et al., 2007; Fénart et al., 2008; Muller et al., 2010). A close phylogenetic relationship between a crop species and a sympatric weed leads to interesting evolutionary dynamics, as ongoing gene flow between the two can shape adaptive evolution of the weed (possibly even through transgene escape), and many evolutionary studies have focused on these related crop-weed systems (eg., Warwick et al., 2003; Morrell et al., 2005; Aono et al., 2006; Campbell et al., 2006). The third mode of weed origination, the niche expansion of wild plants into agroecosystems through plasticity,

adaptation, or preadaptation (when a species requires neither genetic nor phenotypic changes for expansion into new habitats), has received less attention by evolutionary biologists (but see Barrett et al., 1983; Menchari et al., 2007; Welsh and Mohamed, 2011), even though all weeds without close crop relatives must have followed this pathway to agricultural invasion, and this type of weed species is the most common (De Wet and Harlan, 1975). Of the few examples of this mode of agricultural invasion in the literature, the origin of weedy sunflower populations (*Helianthus annuus*) from wild populations is the best documented (Kane and Rieseberg, 2008; Lai et al., 2008).

My study species, waterhemp (*Amaranthus tuberculatus* (Moq.) Sauer), is an herbaceous, outcrossing annual plant native to the Midwestern U.S., where it occurs naturally along riverbanks and in floodplains. Domesticated species of *Amaranthus* are largely absent from its range (Mosyakin and Robertson, 2003). Waterhemp has invaded Midwestern agricultural ecosystems since the 1950s and has become a major problem for farmers since the 1990s (Sauer, 1957; Tranel and Trucco, 2009). In Illinois alone, waterhemp accounts for about 10% of weed control costs for corn and soybean fields, costing farmers an additional $65 million per year (Patrick Tranel, Univ. of IL, pers. comm.). If uncontrolled, it can reduce corn yields by up to 74%, and soybean yields by as much as 56% (Steckel, 2007). As a small-seeded annual with discontinuous germination, waterhemp is a prime example of the class of agricultural weeds that benefited from the widespread adoption of conservation tillage in the late 20th century (Hager et al., 2000; Owen, 2008; Refsell and Hartzler, 2009). Rapid evolution of herbicide resistance has also contributed to waterhemp's success. To date, resistance to five different chemical classes of herbicides has been detected in *A. tuberculatus* populations: PSII-inhibitors (triazines), ALS-inhibitors, HPPD-inhibitors, PPO-inhibitors, and glyphosate (Horak and Peterson, 1995; Foes et

al., 1998; Shoup et al., 2003; Legleiter and Bradley, 2008; Hausman et al., 2011). Furthermore, there is some morphological evidence that the species may have been diverging into two species, one on either side of the Mississippi River (Sauer, 1957), until human disturbance brought the taxa back into contact, and possibly gave rise to the agriculturally invasive form through admixture.

In this study, I used population genetic techniques to test several hypotheses about the origin and evolution of the agricultural weed form of *A. tuberculatus*. I sampled populations of waterhemp across the species' range, and genotyped plants from 38 of these populations using 10 polymorphic microsatellite markers. These population genetic data were used to test the following hypotheses. First, I hypothesized that agriculturally invasive ("crop") populations are genetically differentiated from local native ("non-crop") populations, despite probable high gene flow, because of strong selection for weed-adaptive traits in agricultural environments. Second, I hypothesized (following Sauer, 1957) that *A. tuberculatus* was diverging into two species on opposite sides of the Mississippi River prior to the 20th century, and that the present-day species would retain some genetic and geographical signature of past subdivision into two evolutionary units. The third hypothesis, contingent on the second, was that the agricultural weed originated through hybridization between the two diverged lineages. Based on this last hypothesis, I predicted that populations of waterhemp collected from agricultural fields would show strong evidence of admixture between western and eastern genetic subpopulations.

MATERIALS AND METHODS

Study System

Amaranthus tuberculatus sensu lato (including *A. rudis* sensu Sauer, 1972), is an

herbaceous annual native to North America. The species' range is centered around the Mississippi Valley region, from the Great Plains (roughly as far west as the 100th meridian) eastward to Ohio, and from Louisiana northward to Minnesota, with a northern range boundary in southern Ontario (Figure 2.1). The region of agricultural invasion is more restricted: it is most problematic in the Mississippi Valley region (MO, IL, IA, IN), but also occurs agriculturally in the eastern Great Plains and in parts of Kentucky and Ohio (Tranel and Trucco, 2009). Waterhemp is dioecious (and thus obligately outcrossing) and wind-pollinated, with small one-seeded utricle fruits that may be dehiscent or indehiscent. Natural populations of *A. tuberculatus* are almost always found in disturbed, wet habitats, especially seasonally inundated riverbanks in the Midwest, but also banks of small waterways such as creeks and drainage ditches, lakeshores, and marshy floodplains (Mosyakin and Robertson, 2003). Until Pratt and Clark's 2001 taxonomic study of populations across the species' range, waterhemp was considered two species, distinguished primarily by utricle dehiscence, sepal number, and geographic range: *A. tuberculatus*, the entity with indehiscent utricles almost always found to the east of the Mississippi River; and *A. rudis* (earlier misapplied name = *A. tamariscinus*; see Sauer, 1972), the dehiscent-fruited taxon colloquially understood to be the "weedy" form of waterhemp, found most frequently west of the Mississippi River (Sauer, 1955; 1957; 1972; Figure 2.1). Pratt and Clark found a continuum of morphological and isozyme characters across the range of the more broadly defined *A. tuberculatus*, but some authors still distinguish the two former species as varieties: *A. tuberculatus* var. *tuberculatus* and var. *rudis* (Costea and Tardif, 2003). The latter taxonomy will be used in this dissertation, with Pratt and Clark's species called *A. tuberculatus* sensu lato (s.l.) or simply *A. tuberculatus*.

Sample Collection

I collected 115 populations of *A. tuberculatus* s.l. across the entire species range during field trips in 2009 and 2010. Field trips included: the region around St. Louis in eastern Missouri (multiple trips, July to August, 2009); Nebraska, Kansas, and Oklahoma (September 1-8, 2009); Michigan, Ontario, and Ohio (September 18-29, 2009); Ohio and Indiana (September 16-20, 2010); Illinois and Indiana (September 25-28, 2010); Alabama, Mississippi, Louisiana, and Arkansas (October 15-19, 2010); and central Missouri (October 29, 2011). Populations were located using a combination of herbarium record data and new surveys of typical *A. tuberculatus* habitat along riverbanks, lake shores, and in crop fields. For the areas with agricultural waterhemp populations, both crop field and non-agricultural populations were included in the study. When a population was located, I recorded latitude and longitude coordinates for each population using a Garmin eTrex H handheld GPS unit (Garmin, Olathe, Kansas, USA), and collected a voucher specimen (male and female plants if possible). For each population, either ten dried leaf samples in silica gel were collected, or ten fresh leaf samples were collected, stored in ziploc bags, and kept in a cooler until they could be frozen at -80ºC. The dehiscence of the fruit (considered an important taxonomic character for distinguishing the two varieties within the species) was recorded for each female voucher specimen.

Thirty-eight populations were selected for genotyping to survey the species range (Table 2.1 and Figure 2.2). The St. Louis region was intensively surveyed to determine whether crop and non-crop populations were genetically distinct at a small geographical scale. Ohio was intensively surveyed because it is the edge of the range of agricultural waterhemp, with agricultural populations in ~10 counties west of Columbus but only non-crop populations in the remainder of the state (Figure 2.3).

DNA Extraction and Genotyping

DNA was extracted from each sample with Qiagen DNeasy Plant Mini Kits (Qiagen Inc., Valencia, California, USA). Ten microsatellite loci were amplified and genotyped. Primers, repeat motifs, and sizes of products are listed in Table 2.2. Three of the primer sets are from Lee et al. (2009), and were originally designed from *A. tuberculatus* genomic sequence data. Multiple primer sets from that paper were tested before these three markers were chosen based on consistent amplification and polymorphism. The other seven primer sets were mined from *A. tuberculatus* transcriptomic data using the program SSR Finder (Schroeder, 2003). The transcriptome contigs were provided by Pat Tranel's lab (Univ. IL). These markers were also selected after testing of 14 transcriptome-derived markers in *A. tuberculatus*. In order to multiplex products from different primers in a cost-effective manner, I ordered the forward primers with an M13(-21) sequence (TGTAAAACGACGGCCAGT) at the 5' end, to allow the attachment of a universal fluorescent-dye labeled M13(-21) tag (Schuelke, 2000). The universal tags were labeled with the fluorescent dyes HEX, 6FAM, and NED (Applied Biosystems, Carlsbad, California, USA). In addition, I ordered the reverse primers with a PIG-tail, the sequence "GTTTCTT" at the 5' end of the reverse primer, to facilitate consistent non-template adenylation of the 3' end of the PCR product and to reduce stutter (Brownstein et al., 1996).

PCR was performed on ABI GeneAmp 9700 thermocyclers (Applied Biosystems), in 10 uL reactions containing: 1X GoTaq Flexi Buffer (Promega, Madison, Wisconsin, USA), 2.5 mM MgCl2, 0.05 mM each dNTPs, 0.15 uM M13(-21) dye-labeled tag, 0.04 uM forward primer, 0.16 uM reverse primer, 0.075 uL GoTaq, 3.875 uL nanowater, and 1.25 uL genomic DNA. Amplification conditions were: 94ºC for 5 minutes; then 30 cycles of 94ºC (30 seconds)

denaturation, 51ºC (45 seconds) annealing, 68 ºC (45 seconds) extension (to amplify the product); followed by 8 cycles of 94ºC (30 seconds) denaturation, 48ºC (45 seconds) annealing, 68ºC (45 seconds) extension (to attach the labeled tag); and 72ºC (30 minutes) final extension. PCR products were diluted 1:10 with nanopure water and multiplexed (combining PCR products from up to three loci with different dye labels and different sizes in the same well) with 0.1 uL GeneScan 400HD ROX size standard (Applied Biosystems), denatured for 5 minutes at 95ºC, and genotyped on an ABI Prism 3130x Genetic Analyzer (Applied Biosystems).

Microsatellite data were visualized using GeneMapper v3.7 software (Applied Biosystems). The sizes of alleles at each locus for each individual were recorded by hand and double-checked by repeated amplification and genotyping if more than two peaks appeared (since *A. tuberculatus* is diploid) or unusual allele size classes were observed. If these anomalies were observed twice (which happened very rarely for any particular locus), the data for that marker for that particular individual were coded as missing. Additionally, if genotyping failed for an individual for a particular locus, several subsequent attempts were made to obtain this data before it was coded as missing. The genotyping information was used to create data input files for a variety of population genetic analysis programs, in combination with the geographical coordinates for each population for the spatial genetic programs.

Microsatellite Data Analysis

Microsatellite markers were checked for null alleles using MICRO-CHECKER 2.2.3 (Van Oosterhout et al., 2004). The program Genepop 4.2 (Rousset, 2008) was used to test the probability of Hardy-Weinberg equilibrium for each population (with the Markov chain method to estimate exact p-values), to test for linkage disequilibrium between loci (with Fisher's

method), and to detect private alleles for each population. Popgene 1.31 (Yeh et al. 1997) was used to estimate the average number of observed alleles and effective alleles and average observed and expected heterozygosity over all loci for each population. For all populations combined, the same statistics were estimated for each locus. Weir and Cockerham's theta (an estimate of Fst) was calculated over all loci and all populations using the program FSTAT 2.9.3.2 (Goudet, 1995).

To test for isolation by distance (IBD), I calculated the geographic great circle distance in kilometers between each pair of populations using the Geographic Distance Matrix Generator (Ersts, 2013). I then generated a matrix of pairwise Fst values between populations using Genepop, and combined these two matrices into a data input file for a Mantel test for isolation by distance (with 1000 permutations). This procedure was used to test for IBD across the entire species range, and across subsets of the species range: the Plains states (TR, CHE, TCL, and SaltR populations); Missouri and Illinois; Ohio; and Northern Ohio (OTT, MAU, and PTC populations), Michigan, and Ontario.

To identify the highest-likelihood number of genetic clusters (K) in the data without including geographical information, I applied the program STRUCTURE 2.3.1 (Pritchard et al., 2000). I used the correlated allele frequencies model (on the recommendation of the authors), and used sampling locations (= population assignment) as a prior, which helps with clustering for data with weak genetic STRUCTURE. I ran separate analyses for the admixture and the no admixture ancestry models for both datasets. For the total species range dataset, I ran the analysis for K=1 to K=10; for the Ohio dataset, from K=1 to K=6; and for the St. Louis dataset, from K=1 to K=5, with three runs per K, 100,000 Markov Chain Monte Carlo (MCMC) burnin steps, and 500,000 MCMC steps after the burnin for all datasets. The separate analyses of the

Ohio and St. Louis datasets were conducted to examine fine-scale structure in these parts of the range that were intensively sampled. To estimate the number of genetic clusters from the ln Probability (X|K) values output by STRUCTURE, I used the delta K method of Evanno et al. (2005).

AMOVA was employed in GenAlEx 6.5 (Peakall and Smouse, 2012) to test several hypotheses about the partitioning of genetic variance. Within-individual estimates were suppressed, and 999 permutations were used to generate a range of F-statistics under the null hypothesis of no genetic subdivision within the total dataset. Two hypotheses were tested regarding population differentiation. With the total species range dataset (38 populations, Figure 2.2), the sampled populations were hypothesized to belong to two genetic clusters differentiated according to whether a population occurred in a crop field or a non-crop environment (Table 2.1), or two genetic clusters based on STRUCTURE results (see below). Admixed populations from the STRUCTURE analysis were assigned to a single genetic cluster for the AMOVA analyses based on their predominant cluster affinity (i.e., predominant membership assignment). The same hypotheses were tested for the limited dataset of the Ohio populations (12 populations, Figure 2.3). Finally, the St. Louis populations (AAF, WSR, GTP, WSS, and EMN) were tested for genetic subdivision by crop vs. non-crop environment. Ohio and St. Louis were chosen for these analyses because of the intensive sampling of both types of habitats in these regions.

Two different programs were used to include spatial information in the estimation of genetic clusters for the total species range dataset, with admixture models rather than no admixture models as recommended by François and Durand (2010). The first program was TESS 2.3 (Durand et al. 2009), a Bayesian clustering algorithm which has two priors: a spatial neighborhood network (based on the Voroni tessellation) and a Markov Hidden Gaussian

Random Field model (when the admixture model is employed). Just like STRUCTURE, it can subsequently be used to identify the highest-likehood number of genetic clusters and the assignment of each individual and population to these clusters, and it gives deviance information criterion (DIC) values rather than ln Probability (X|K) values as an estimate of the likelihood of each K value. I ran the program for the total species' range dataset with the CAR model of admixture, using 50,000 total sweeps (10,000 of these burnin). I ran the analysis for K=2 to K=10, with three runs per K. The graphical output from TESS includes a bar graph showing genetic assignment of individuals in different colors, and also a Voroni tessellation diagram showing the spatial genetic assignment of populations in the same colors.

The second spatial clustering program was BAPS v5.3, Bayesian Analysis of Population Structure (Corander et al., 2003). Unlike STRUCTURE and TESS, BAPS does not use MCMC to infer K. Instead, BAPS uses a stochastic search algorithm that considers multiple K values simulataneously to directly estimate the number of genetic clusters and assign individuals to those clusters using mixture analysis. Geographic localities of populations can be employed as priors, using the "spatial clustering of groups" option (Corander et al. 2008). For the mixture analysis, the user must specify the maximum number of clusters expected in the group, and several K_{max} values may be specified. I set K_{max} = 2, 5, 10, and 20, with three runs per K. BAPS analyzes admixture using the mixture analysis results from the highest-probability K as input. For this analysis, I used a minimum population size of five, 200 iterations, 100 reference individuals for each population, and 20 iterations of reference individuals (as suggested by the manual). The graphical output from BAPS is an "admixture partition" bar graph showing genetic assignment of individuals, and a Voroni tessellation diagram showing the spatial genetic assignment of populations.

RESULTS

Genetic Diversity and Isolation by Distance

Populations largely conformed to Hardy-Weinberg expectations. MICRO-CHECKER revealed that none of the loci were consistently more homozygous than expected, and therefore there was no evidence for null alleles in the dataset. Likewise, no single population showed a deviation from Hardy-Weinberg equilibrium at more than one locus. Genepop Hardy-Weinberg probability tests gave slightly different values, showing that the locus AAC1 was out of Hardy-Weinberg equilibrium for five populations, and the population MAU was out of Hardy-Weinberg equilibrium when all loci were taken into account ($p=0.0236$). STRUCTURE analyses were run with and without this locus and population, with no change in the best-supported K-value and very little change in the bar graph (results not shown).

Genepop showed no linkage disequilibrium between loci. Averages for the observed number of alleles, effective number of alleles, and observed and expected heterozygosity for each population are shown in Table 2.3, and the same statistics are shown for each locus over all populations in Table 2.4. Expected heterozygosity ranged between 0.4245 and 0.6829 for individual populations, with a mean of 0.5557 over all populations, indicating high within-population genetic diversity. Populations in the western half of the species range tended to have higher average observed and effective numbers of alleles than populations in the eastern half of the species range. Seven populations in Indiana and Ohio and one population in Illinois (PEK) had higher average observed than expected heterozygosity, potentially suggesting recent admixture (Table 2.3). The number of alleles per locus ranged from 6 to 20, and in general both observed and expected heterozygosity per locus were high, with the exception of the locus

ATC9, which had approximately one effective allele (Table 2.4). Weir and Cockerham's theta was 0.075 over all loci, with a range of 0.029-0.186 for individual loci, showing overall low genetic differentiation between populations.

Mantel tests performed in Genepop showed isolation by distance across the entire species range ($p<0.00001$, Figure 2.4). Pairwise Fst values between populations ranged from 0.0013 to 0.2681. For subsets of the species range, there was no isolation by distance at the state or bi-state level, or across the three Plains states. However, the dataset composed of populations from northern Ohio, Michigan, and Ontario did show weak isolation by distance ($p=0.037$).

Population Structure

The STRUCTURE plots of delta K (the second order rate of change of K) as a function of K are shown in Figures 2.5, 2.6, and 2.7, for the total species range dataset, the Ohio dataset, and the St. Louis dataset, respectively. The highest value of delta K or the value at which delta K plateaus indicates the inferred K value in the data. For the total species range dataset, there were two genetic clusters, one characteristic of the western part of the geographic range and one characteristic of the eastern part, with substantial admixture inferred for the populations PEK (IL), KANK, WAB, IND (IN), and the OH agricultural waterhemp region populations (Figures 2.8 and 2.9). For the Ohio dataset, there were also two genetic clusters (albeit more weakly supported by delta K), one cluster in the agricultural waterhemp region of Ohio and one in the Ohio River region, with admixture between the two clusters inferred for populations in northern Ohio and in the southern Ohio population BTL (Figure 2.10). On the other hand, the St. Louis dataset supported the presence of only one genetic cluster: STRUCTURE and the delta K method are not designed to detect a single genetic cluster, but the bar plots for all K values above one

equally divide every individual between the K genetic clusters, which is an indication that no genetic structure exists (Thinglum, 2010; see Figure 2.11 for an example at K=2). Analyzing the same datasets with the no-admixture model yielded the same inferred numbers of clusters (results not shown).

Results of the AMOVA analyses are shown in Table 2.5. No clear differentiation between agricultural and non-agricultural waterhemp was observed at any geographical scale. When the total species range dataset was divided into crop and non-crop regions, 7% of molecular variance was partitioned between populations (df=36, p=0.001), but none of the molecular variance was partitioned between regions (df=1, p=0.232). When the dataset was subdivided according to genetic clusters from the STRUCTURE K=2 results (see below), 5% of molecular variance was partitioned between populations (df=36, p=0.001) and 5% was partitioned between regions (df = 1, p=0.001). For the Ohio dataset, a division of crop vs. non-crop regions yielded 6% of variation between populations (df=10, p=0.001) and 1% between regions (df=1, p=0.001), whereas when the dataset was subdivided according to the STRUCTURE K=2 results, 5% of variation was found between populations (df=10, p=0.001) and 3% between regions (df=1, p=0.001). For the St. Louis dataset, none of the molecular variance was partitioned between crop and non-crop regions (df=1, p=0.727), and 4% of the variance was partitioned between populations (df=3, p=0.001). Together, these patterns suggest no differentiation between crop and non-crop populations.

For the spatial genetic analysis of the total species range dataset with TESS, the DIC values are plotted in Figure 2.12. The lowest DIC value should indicate the number of inferred genetic clusters, but because K=1 cannot be estimated by TESS, it is difficult to detect a plateau starting at K=2. However, the bar charts assign individuals to clusters consistently for all three

runs for K=2, while for K=3 and above, the cluster assignment is similar to K=2 with very minor increases in the inferred number of clusters for a few individuals (results not shown). Also, the "hard clustering" Voroni tesselation diagrams show two genetic clusters for K=3 and above, with population assignments nearly identical to those for the three (identical) runs of K=2 (results not shown). These patterns suggest that, like STRUCTURE, TESS supports the existence of two genetic clusters in the data. In the TESS bar graph output, population and individual assignments to the two clusters are quite similar to assignments from STRUCTURE to the comparable groups (with the western genetic cluster shown in red and the eastern cluster in green for both programs), except that the KNK population (IL) shows more admixture and the northern Ohio populations show slightly less in the TESS analysis (Figure 2.13). Each population is assigned to a single cluster (with no admixture shown) in the "hard clustering" Voroni tessellation diagram: the populations west of the IND population (Indianapolis, IN), are almost all assigned to one genetic cluster, and the populations to the east of IND are almost all assigned to the other. The exceptions are the western population PEK (IL), which is assigned to the eastern cluster, and the eastern populations STW and CAN (OH), which are assigned to the western cluster (Figure 2.14). The agreement between the STRUCTURE and TESS results suggests that a genetic signature of the two previously-diverging lineages in the species still remains in the present-day species.

For the spatial analysis of the total species range dataset with BAPS, the program identified the highest K value as K=3 (for K_{max} = 5, 10, and 20). Population and individual assignments to these three clusters are shown in the admixture partition bar graph (resulting from a mixture analysis, followed by admixture analysis) (Figure 2.15). The population assignments largely correspond to the same western/eastern divide seen in the STRUCTURE and TESS

analyses, with the dividing line between VIGO (IN) and KANK (IN), PEK assigned to the eastern cluster, and IND, STW, and SCIO assigned to the western cluster. The MC population (OH) was the only population in the third cluster. BAPS identified only eight individuals as exhibiting admixture, but these individuals were in the central populations (PEK, KNK, KANK, IND) and agricultural waterhemp region of Ohio (STW, SCIO), which correspond to populations with high admixture in STRUCTURE and TESS. The Voroni tessellation diagram shows the same population genetic assignments spatially (without admixture) (Figure 2.16). These results provide further support for the hypothesis of two genetic lineages within the species that recently came back into contact.

DISCUSSION

The combined results from AMOVA, isolation by distance tests, STRUCTURE, TESS, and BAPS are largely congruent and paint an interesting picture of the recent origin and evolution of the agricultural weed form of *Amaranthus tuberculatus*. First, there is no evidence to support the hypothesis of genetic differentiation at neutral markers between "crop" and "non-crop" populations of waterhemp, either over the entire species range or at a smaller geographic scale. AMOVAs show no partitioning of genetic variance between regions when the populations are divided by agricultural vs. non-agricultural habitat, either for the total species range or the St. Louis area, and STRUCTURE revealed a single genetic cluster in the five St. Louis populations. The Ohio populations show a significant amount of variation between regions with AMOVA (1%) when the agricultural vs. non-agricultural division is applied, but this is probably because four out of six populations in the "agricultural waterhemp region" of Ohio (which corresponds to a genetic cluster found in the STRUCTURE analysis) are crop populations.

The isolation by distance analyses, which show no isolation by distance at scales smaller than three states or provinces, and the low overall Fst in the total range dataset also support the idea that gene flow homogenizes neutral genetic variation over large areas of the species range, overwhelming the effects of selective sweeps on functional genes due to changing management practices in agricultural environments (Thinglum, 2010). Further support for this idea comes from the species' biology, as it is obligately outcrossing, wind pollinated, and probably has a very large effective population size and very high effective recombination across the genome (Thinglum et al., 2011). Follow-up work using denser marker coverage could potentially reveal the specific genomic regions showing adaptive differentiation between agricultural and natural environments (e.g., Loh et al., 2008; Bouchet et al., 2012). Differentiation between nearby wild and weedy populations has been detected using 106 microsatellite markers in sunflowers (Kane and Rieseberg, 2008; Lai et al, 2008); however, a European study of the rapidly expanding, wind-pollinated weed blackgrass (*Alopecurus myosuroides*) found that even though agricultural populations experienced strong selection from herbicide application, this did not modify their genetic structure at 116 AFLPs distributed across the genome (Menchari et al., 2007).

The second hypothesis, that the species was formerly diverging into two evolutionary units, was supported by my data. STRUCTURE and TESS both recovered two genetic clusters from the total species range dataset, and at the range edges, the geographical structure of these clusters corresponds closely to the hypothesized eastern/western divide between the two former taxonomic units (Sauer, 1957) (Figure 2.1). BAPS recovered an additional cluster consisting of one Ohio agricultural population, which might have distinctive multigene allele frequencies due to admixture. It appears from my data as though the Mississippi River is no longer the geographical divide between the two genetic clusters; instead, the western genetic cluster extends

into Indiana. Interestingly this boundary shift was documented more than half a century ago by Sauer (1957), who observed from herbarium specimen records that the western taxon, now called *A. tuberculatus* var. *rudis*, had been moving steadily northward and eastward across the Mississippi River since the 1850s, into the range of *A. tuberculatus* var. *tuberculatus*. Furthermore, he noted that this movement was associated with agricultural invasion: the earliest records of *A. tuberculatus var. rudis* in Illinois (1940s) and Indiana (1950s) are reports from agricultural fields (Sauer, 1957).

Pratt and Clark's (2001) analysis of 27 morphological characters and 14 isozyme loci across the range of *A. tuberculatus* s.l. revealed a continuum of morphological character states and isozyme alleles across the entire range. On the basis of no clear clustering in a PCA of these characters, they declared the two taxa to be one variable species. The observed continuum is not surprising, given the geographical overlap between the two varieties that has occurred in the middle of the range as *A. tuberculatus* var. *rudis* pushed eastward. In my own voucher specimens, the morphological character of utricle dehiscence is nearly constant at the western and eastern ends of the range (the Plains states and Ontario), and extremely variable both within and among populations in the range center (Table 2.1). Both Sauer and Pratt and Clark were primarily focused on taxonomy, and tended not to focus on potentially interesting population-level patterns. My application of a relatively recent genetic tool, microsatellite genotyping, has largely confirmed their broad-scale observations and has also shed more light on the origins of the agricultural weed form.

Spatial genetic clustering allows the use of information beyond genotype data (such as spatial autocorrelation and geographical trends) in inferring population structure, and can be especially useful when closely-related taxa come into secondary contact at regional geographic

scales (François and Durand, 2010). The differences in clustering between the spatial genetic programs TESS and BAPS are probably due to differences in the assumptions of the underlying Bayesians clustering methods. While TESS uses Markov methods (similar to STRUCTURE) to find the highest likelihood for each K value independently, BAPS uses a stochastic search algorithm that directly estimates the most likely K value. Furthermore, BAPS estimates admixture after partitioning the data into clusters with a mixture model, while TESS estimates admixture and the likelihood of each K value simultaneously. Because of these differences, the discovery by both programs of two major genetic clusters in my dataset is strong support for this result.

The last hypothesis, that weedy waterhemp was created through hybridization between the two evolutionary units in *A. tuberculatus* s.l., was not supported by my data. If this were the case, one would expect that the agricultural populations of waterhemp would show strong evidence of admixture between the two genetic clusters. Instead, almost all of the Missouri, Illinois, and Indiana agricultural populations show a very strong affinity with the western (red) genetic cluster in STRUCTURE, TESS, and BAPS, with one exception (the Illinois population KNK is strongly admixed in TESS and admixed in BAPS). At the edge of the range of agricultural waterhemp in Ohio, admixture is prevalent in all populations (including riverbank populations) within the agricultural waterhemp region according to STRUCTURE analysis of the total range dataset. However, when only the Ohio populations are analyzed in STRUCTURE, admixture was quite low in the four agricultural populations (CAN, GTB, MC, and RT29), which were mainly assigned to the western genetic cluster.

The geographical pattern of admixture in the data suggests that the movement of *A. tuberculatus* var. *rudis* eastward almost completely replaced *A. tuberculatus* var. *tuberculatus*

populations in natural environments in Illinois (e.g., the natural populations RIP and KEY have very little signature of the eastern genetic cluster). However, the more northern Illinois population PEK shows strong admixture in every analysis, and the northern population KNK shows strong admixture in two analyses, which suggests that the invasion of *A. tuberculatus* var. *rudis* may have a northern geographical boundary. The same boundary is apparent in Indiana, where the two northern populations KANK and WAB either belong predominantly to the eastern genetic cluster or are strongly admixed, depending on the analysis. The middle of Indiana is almost entirely the western genetic cluster on the very western edge of the state (VIGO), but admixed in the center (IND), and almost entirely the eastern genetic cluster along the Ohio River in the south (AUR), suggesting a southern boundary for the invasion as well.

The more extensive sampling of Ohio also supports the idea that the invasion of *A. tuberculatus* var. *rudis* is confined to the "agricultural waterhemp region" in the middle of the state (labeled in Figures 2.8, 2.10, and 2.13). The populations in this region are strongly admixed or western genetically (TESS and STRUCTURE bar graphs), and several populations are primarily the western genetic cluster according to the "hard clustering" TESS and BAPS Voroni diagrams (Figures 2.14 and 2.16). The southern populations along the Ohio River are almost entirely the eastern genetic cluster, and the southern population BTL and the northern populations OTT, MAU, and PTC are eastern or strongly admixed according to different analyses (with only STRUCTURE supporting admixture). The range boundary of agricultural waterhemp is around Columbus, Ohio, and natural populations were not sampled in the eastern half of the state. The inclusion of more eastern populations could confirm the idea that *A. tuberculatus* var. *rudis* genetic material also hits a range boundary in western Ohio.

Altogether, these patterns of genetic clustering point to a geographical invasion of *A.*

tuberculatus var. *rudis* almost directly eastward through the primary agricultural regions of the eastern states, facilitated by introduction first in crop fields (as observed by Sauer, 1957). Waterhemp weed seeds are extensively moved around by farm equipment, which is often shared between farms and transported long distances (Patrick Tranel, University of IL, pers. comm.). With the evolution of resistance to multiple herbicide classes in the species, the spread of *A. tuberculatus* var. *rudis* throughout the Midwest became practically inevitable. The reasons for the northern, southern, and eastern geographical boundaries deserve further study: the Ohio boundary may involve soil substrate (which changes abruptly in the middle of the state), and the northern and southern boundaries are more likely to involve differences in climate and topography.

Given the wind-dispersed pollen and obligately outcrossing nature of waterhemp, it is perhaps surprising that any genetic signature of the two subspecies, let alone a genetic signature of the eastward invasion, still exists. Pollen of *A. tuberculatus* is viable for up to 120 hours, allowing for long-distance dispersal, although most pollen fertilizes plants within 50 meters in field trials (Liu et al., 2012). Furthermore, there is no evidence for pre- or postzygotic reproductive barriers between the two varieties (Murray 1940). Interestingly, the genetic pattern of invasion closely corresponds to predictions by Currat et al. (2008), who modeled introgression between an invader and a compatible local species using coalescent simulations. They found that introgression of neutral genes happens extensively unless strong reproductive or geographic barriers exist, and that gene flow is almost entirely from the local species to the invader at the invasion front. This is shown in my results by the greater signature of admixture in the Ohio agricultural waterhemp region (the range edge) than in the invaded regions of Illinois, where *A. tuberculatus* var. *rudis* has had more time to build up population sizes in agricultural fields and

entirely swamp out the eastern subspecies in natural habitats.

As stated above, hybridization between the two varieties of *A. tuberculatus* does not appear to have been involved in the formation of the weed form. It is possible that introgression of weediness alleles from another *Amaranthus* species led to the evolution of weediness in *A. tuberculatus* var. *rudis*, as previously hypothesized by Tranel et al. (2002). My study was not designed to test this hypothesis, but a STRUCTURE analysis of eight populations of the sympatric monoecious species *A. hybridus* and nearby *A. tuberculatus* populations with seven microsatellite markers showed no evidence of introgression between the two species (data not shown). However, this analysis probably would not have identified any potential adaptive introgression between the species, given that the genetic regions involved in agricultural invasion are unknown in *Amaranthus,* and it is unlikely that they are tightly linked to these neutral markers. A very small number of genetic regions from *A. hybridus* could have conferred weediness in *A. tuberculatus*, and these regions could have been quickly disassociated from other *A. hybridus* genes through extensive recombination during backcrossing with *A. tuberculatus.* Hybrids between *A. hybridus* and *A. tuberculatus* are frequently identified in the field based on morphology (Pratt, 1999). However, at least for the trait of herbicide resistance, agriculturally adaptive alleles in *A. tuberculatus* are not derived from introgression between the two species: a greenhouse experiment by the Tranel lab showed unidirectional transfer of alleles (including herbicide resistance alleles) from *A. hybridus* to *A. tuberculatus* (Trucco et al. 2009), and herbicide resistance at the ALS locus appears to be evolving independently in the two species (Tranel and Trucco, 2009).

Returning to the major question of the chapter — how the weedy form of *A. tuberculatus* arose — introgression between the two varieties is not supported as a causative factor in this

study, and there is no evidence about the role of hybridization with another weedy *Amaranthus* species one way or the other. The most likely scenario is that the weed form is simply *A. tuberculatus* var. *rudis*, which was preadapted to invade agricultural environments. When Mississippi Valley environments became increasingly dominated by agriculture in the 20th century, due to large-scale mechanized farming and the channeling of rivers for the greater agricultural availability of floodplain habitats (Ghersa et al., 1994), *A. tuberculatus* var. *rudis* was already well-suited to coexist and compete with crops in these new environments. Later in the 20th century, the further expansion of waterhemp as a weed was facilitated by the widespread adoption of no-till agriculture and herbicide-based weed control (Costea et al., 2005). The idea that *A. tuberculatus* var. *rudis* was already "weedy" and might not have required genetic changes to be successful in agricultural ecosystems is supported by Sauer's description of the taxon, in which he states that in contrast to *A. tuberculatus* var. *tuberculatus*, var. *rudis* has "very definite weedy tendencies," and one-third of the herbarium collections of the species are from artificial, anthropogenically-disturbed habitats (Sauer, 1955).

Preadaptation is not the same concept as plasticity (i.e., Baker's (1965) "general-purpose genotype"), as both imply that a species does not undergo a genetic adaptation in response to selection, but preadaptation suggests that a species might not change its phenotype either. Evidence for local adaptation as well as preadaptation of agriculturally-invasive *A. tuberculatus* is presented in Chapter 3, but plasticity in response to varying environments has not been studied in the species. This is an intriguing avenue for future studies of waterhemp, given that plasticity is often hypothesized to be very important for invasive plants with little genetic variation (e.g., Parker et al., 2003), but its role in invasive species with high genetic variation has seldom been examined, despite the potential for evolution of plasticity itself in these species (but see Sexton et

al., 2002). There is some evidence for preadaptation in plant invasions from the invasive-species literature (Kolar and Lodge, 2001; Schlaepfer et al., 2010; Van Kleunen et al., 2011).

In this study, I have built on the observations of Sauer (1957) and Pratt and Clark (2001) to present a new hypothesis about the origin of the agricultural weed form of *A. tuberculatus*. Evolution in response to agricultural management practices is ongoing in this species, as exemplified by its continual adaptation to new herbicides (e.g., Hausman et al., 2011), and *A. tuberculatus*' current range boundaries may shift in response to evolution or land use changes. This research shows that agricultural weeds unrelated to domesticated plants can have great potential as evolutionary model systems.

LITERATURE CITED

AONO, M., S. WAKIYAMA, M. NAGATSU, N. NAKAJIMA, M. TAMAOKI, A. KUBO, AND H. SAJI. 2006. Detection of feral transgenic oilseed rape with multiple-herbicide resistance in Japan. *Environmental Biosafety Research* 5: 77-87.

BAKER, H. G. 1965. Characteristics and modes of origin of weeds. *In* H. G. Baker and G. L. Stebbins, [eds.], The genetics of colonizing species. Academic Press, New York, USA, pgs. 147-168.

BAKER, H. G. 1974. The evolution of weeds. *Annual Review of Ecology and Systematics* 5: 1-24.

BARRETT, S. C. H. 1983. Crop mimicry in weeds. *Economic Botany* 37: 255-282.

BOUCHET, S., D. POT, M. DEU, J.-F. RAMI, C. BILLOT, X. PERRIER, R. RIVALLAN, ET AL. 2012. Genetic structure, linkage disequilibrium and signature of selection in sorghum: lessons from physically anchored DArT markers. PLoS ONE 7: e33470.

BROWNSTEIN, M. J., J. D. CARPTEN, AND J. R. SMITH. 1996. Modulation of non-templated nucleotide addition by *Taq* DNA polymerase: primer modifications that facilitate genotyping. *BioTechniques* 20: 1004-1010.

BURGER, J. C., S. LEE, AND N. C. ELLSTRAND. 2006. Origin and genetic structure of feral rye in the western United States. *Molecular Ecology* 15: 2527-2539.

CAMPBELL, L. G., A. A. SNOW, AND C. E. RIDLEY. 2006. Weed evolution after crop gene introgression: greater survival and fecundity of hybrids in a new environment. *Ecology Letters* 9: 1198-1209.

CORANDER, J., P. WALDMANN, AND M. J. SILLANPÄÄ. 2003. Bayesian analysis of genetic differentiation between populations. *Genetics* 163: 367-374.

CORANDER, J., J. SIRÉN, AND E. ARJAS. 2008. Bayesian spatial modeling of genetic population structure. *Computational Statistics* 23: 111-129.

COSTEA, M., AND F. J. TARDIF. 2003. Conspectus and notes on the genus *Amaranthus* in Canada. *Rhodora* 105: 260-281.

COSTEA, M., S. E. WEAVER, AND F. TARDIF. 2005. The biology of invasive alien plants in Canada. 3. *Amaranthus tuberculatus* (Moq.) Sauer var. *rudis* (Sauer) Costea & Tardif. *Canadian Journal of Plant Science* 85: 507-522.

CURRAT, M., M. RUEDI, R. J. PETIT, AND L. EXCOFFIER. 2008. The hidden side of invasions: massive introgression by local genes. *Evolution* 62: 1908-1920.

DE WET, J. M. J., AND J. R. HARLAN. 1975. Weeds and domesticates: evolution in the man-made habitat. *Economic Botany* 29: 99-107.

DURAND, E., F. JAY, O. E. GAGGIOTTI, AND O. FRANÇOIS. 2009. Spatial inference of admixture proportions and secondary contact zones. *Molecular Biology and Evolution* 26: 1963-1973.

ERSTS, P.J. [Internet] Geographic Distance Matrix Generator (version 1.2.3). American Museum of Natural History, Center for Biodiversity and Conservation. Available from http://biodiversityinformatics.amnh.org/open_source/gdmg. Accessed on 2013-6-2.

EVANNO, G., S. REGNAUT, AND J. GOUDET. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14: 2611-2620.

FÉNART, S., J-F. ARNAUD, I. DE CAUWER, AND J. CUGUEN. 2008. Nuclear and cytoplasmic genetic diversity in weed beet and sugar beet accessions compared to wild relatives: new insights into the genetic relationships within the *Beta vulgaris* complex species. *Theoretical and Applied Genetics* 116: 1063-1077.

FOES, M. J., P. J. TRANEL, L. M. WAX, AND E. W. STOLLER. 1998. A biotype of common waterhemp (*Amaranthus rudis*) resistant to triazine and ALS herbicides. *Weed Science* 46: 514-520.

FRANÇOIS, O., AND E. DURAND. 2010. Spatially explicity Bayesian clustering models in population genetics. *Molecular Ecology Resources* 10: 773-784.

GHERSA, C. M., M. L. ROUSH, S. R. RADOSEVICH, AND S. M. CORDRAY. 1994. Coevolution of agroecosystems and weed management. *BioScience* 44: 85-94.

GOUDET, J. 1995. FSTAT Version 1.2: a computer program to calculate F-statistics. *Journal of Heredity* 86: 485-486. FSTAT 2.9.3.2 software available from: http://www2.unil.ch/popgen/softwares/fstat.htm

HAGER, A., L. WAX, W. SIMMONS, AND C. SPRAGUE. 2000. Waterhemp management in Illinois agronomic crops. *In* the University of Illinois Extension [eds.], 2000 Illinois Agricultural Pest Management Handbook, 91-100. The University of Illinois, Urbana-Champaign, Illinois, USA.

HAUSMAN, N. E., S. SINGH, P. J. TRANEL, D. E. RIECHERS, S. S. KAUNDUN, N. D. POLGE, D. A. THOMAS, AND A. G. HAGER. 2011. Resistance to HPPD-inhibiting herbicides in a population of waterhemp (*Amaranthus tuberculatus*) from Illinois, United States. *Pest Management Science* 67: 258-261.

HILGENFELD, K. L., A. R. MARTIN, D. A. MORTENSEN, AND S. C. MASON. 2004. Weed management in a glyphosate resistant soybean system: weed species shifts. *Weed Technology* 18: 284-291.

HORAK, M. J., AND D. E. PETERSON. 1995. Biotypes of Palmer amaranth (*Amaranthus palmeri*) and common waterhemp (*Amaranthus rudis*) are resistant to imazethapyr and thifensulfuron. *Weed Technology* 9: 192-195.

LEGLEITER, T. R., AND K. W. BRADLEY. 2008. Glyphosate and multiple herbicide resistance in common waterhemp (*Amaranthus rudis*) populations from Missouri. *Weed Science* 56: 582-587.

KANE, N. C., AND L. H. RIESEBERG. 2008. Genetics and evolution of weedy *Helianthus annuus* populations: adaptation of an agricultural weed. *Molecular Ecology* 17: 384-394.

KOLAR, C. S., AND D. M. LODGE. 2001. Progress in invasion biology: predicting invaders. *Trends in Ecology and Evolution* 16: 199-204.

LAI, Z., N. C. KANE, Y. ZOU, AND L. H. RIESEBERG. 2008. Natural variation in gene expression between wild and weedy populations of *Helianthus annuus*. *Genetics* 179: 1881-1890.

LOH, Y.-H. E., L. S. KATZ, M. C. MIMS, T. D. KOCHER, S. V. YI, AND J. T. STREELMAN. 2008. Comparative analysis reveals signatures of differentiation amid genomic polymorphism in Lake Malawi cichlids. *Genome Biology* 9: R113.

LONDO, J. P., AND B. A. SCHAAL. 2007. Origins and population genetics of weedy red rice in the USA. *Molecular Ecology* 16: 4523-4535.

LIU, J., A. S. DAVIS, AND P. J. TRANEL. 2012. Pollen biology and dispersal dynamics in waterhemp (*Amaranthus tuberculatus*). *Weed Science* 60: 416-422.

MENCHARI, Y., C. DÉLYE, AND V. LE CORRE. 2007. Genetic variation and population structure in black-grass (*Alopecurus myosuroides* Huds.), a successful, herbicide-resistant, annual grass weed of winter cereal fields. *Molecular Ecology* 16: 3161-3172.

MORRELL, P. L., T. D. WILLIAMS-COPLIN, A. L. LATTU, J. E. BOWERS, J. M. CHANDLER, AND A. H. PATERSON. 2005. Crop-to-weed introgression has impacted allelic composition of johnsongrass populations with and without recent exposure to cultivated sorghum. *Molecular Ecology* 14: 2143-2154.

MOSYAKIN, S., AND K. R. ROBERTSON. 2003. *Amaranthus*. *In* Flora of North America Editorial Committee [eds.], Flora of North America North of Mexico, Vol. 4. Magnoliophyta: Caryophyllidae, Pt. 1. Oxford University Press, New York, New York, USA.

MULLER, M.-H., M. LATREILLE, AND C. TOLLON. 2010. The origin and evolution of a recent agricultural weed: population genetic diversity of weedy populations of sunflower (*Helianthus annuus* L.) in Spain and France. *Evolutionary Applications* 4: 499-514.

MURRAY, J.M. 1940. The genetics of sex determination in the family Amaranthaceae. *Genetics* 25: 409–431.

OLSEN, K. M., A. L. CAICEDO, AND Y. JIA. 2007. Evolutionary genomics of weedy rice in the USA. *Journal of Integrative Plant Biology* 49: 811-816.

OWEN, M. D. K. 2008. Weed species shifts in glyphosate-resistant crops. *Pest Management Science* 64: 377-387.

PARKER, I. M., J. RODRIGUEZ, AND M. E. LOIK. 2003. An evolutionary approach to understanding the biology of invasions: local adaptation and general-purpose genotypes in the weed Verbascum thapsus. *Conservation Biology* 17: 1-15.

PEAKALL, R., AND P. E. SMOUSE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. *Bioinformatics* 28: 2537-2539.

PRATT, D. B. 1999. Taxonomic revision of *Amaranthus rudis* and *A. tuberculatus* (Amaranthaceae). Masters Thesis, Iowa State University.

PRATT, D. B., AND L. G. CLARK. 2001. *Amaranthus rudis* and *A. tuberculatus*, one species or two? *Journal of the Torrey Botanical Society* 128: 282-296.

PRITCHARD, J. K., M. STEPHENS, AND P. DONNELLY. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.

REFSELL, D. E., AND R. G. HARTZLER. 2009. Effect of tillage on common waterhemp (*Amaranthus rudis*) emergence and vertical distribution of seed in the soil. *Weed Technology* 23: 129-133.

ROUSSET, F. 2008. Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources* 8: 103-106.

SAUER, J. D. 1955. Revision of the dioecious amaranths. *Madroño* 13: 5-46.

SAUER, J. D. 1957. Recent migration and evolution of the dioecious amaranths. *Evolution* 11: 11-31.

SAUER, J. D. 1972. The dioecious amaranths: a new species name and major range extensions. *Madroño* 21: 427-434.

SCHLAEPFER, D. R., M. GLÄTTLI, M. FISCHER, AND M. VAN KLEUNEN. 2010. A multi-species experiment in their native range indicates pre-adaptation of invasive alien plant species. *New Phytologist* 185: 1087-1099.

SCHROEDER, S. 2003. SSR Finder. Maize Mapping Project, University of Missouri, Columbia, MO, USA.

SCHUELKE, M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18: 233-234.

SEXTON, J. P., J. K. MCKAY, AND A. SALA. 2002. Plasticity and genetic diversity may allow saltcedar to invade cold climates in North America. *Ecological Applications* 12: 1652-1660.

SHOUP, D. E., K. AL-KHATIB, AND D. E. PETERSON. 2003. Common waterhemp (*Amaranthus rudis)* resistance to protoporphyrinogen oxidase-inhibiting herbicides. *Weed Science* 51: 145-150.

STECKEL, L. E. 2007. The dioecious *Amaranthus* spp.: here to stay. *Weed Technology* 21: 567-570.

THINGLUM, K. A. 2010. Population genetics of herbicide resistance in waterhemp. Masters Thesis, University of Illinois at Urbana-Champaign. Available through the Illinois Digital Environment for Access to Learning and Scholarship.

THINGLUM, K. A., C. W. RIGGINS, A. S. DAVIS, K. W. BRADLEY, K. AL-KHATIB, AND P. J. TRANEL. 2011. Wide distribution of the waterhemp (*Amaranthus tuberculatus*) ΔG210 *PPX2* mutation, which confers resistance to PPO-inhibiting herbicides. *Weed Science* 59: 22-27.

TRANEL, P. J., AND F. TRUCCO. 2009. 21st century weed science: a call for *Amaranthus* genomics. *In* C. N. Stewart Jr. [ed.], Weedy and Invasive Plant Genomics, 53-81. Blackwell Publishing, Ames, IA, USA.

TRANEL, P. J., J. J. WASSOM, M. R. JESCHKE, AND A. L. RAYBURN. 2002. Transmission of herbicide resistance from a monoecious to a dioecious weedy *Amaranthus* species. *Theoretical and Applied Genetics* 105: 674-679.

TRUCCO, F., T. TATUM, A. L. RAYBURN, AND P. J. TRANEL. 2009. Out of the swamp: unidirectional hybridization with weedy species may explain the prevalence of *Amaranthus tuberculatus* as a weed. *New Phytologist* 184: 819-827.

VAN KLEUNEN, M., D. R. SCHLAEPFER, M. GLAETTLI, AND M. FISCHER. 2011. Preadapted for invasiveness: do species traits or their plastic response to shading differ between invasive and non-invasive plant species in their native range? *Journal of Biogeography* 38: 1294-1304.

VAN OOSTERHOUT, C., W. F. HUTCHINSON, D. P. M. WILLS, AND P. SHIPLEY. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4: 535-538.

VIGUEIRA, C. C., K. M. OLSEN, AND A. L. CAICEDO. 2013. The red queen in the corn: agricultural weeds as models of rapid adaptive evolution. Heredity 110: 303-311.

WARWICK, S. I., M.-J. SIMARD, A. LÉGÈRE, H. J. BECKIE, L. BRAUN, B. ZHU, P. MASON, G. SÉGUIN-SWARTZ, AND C. N. STEWART. 2003. Hybridization between transgenic *Brassica napus* L. and its wild relatives: *Brassica rapa* L., *Raphanus raphanistrum* L., *Sinapis arvensis* L., and *Erucastrum gallicum* (Willd.) O.E. Schulz. *Theoretical and Applied Genetics* 107: 528-539.

WELSH, A. B., AND K. I. MOHAMED. 2011. Genetic diversity of *Striga hermonthica* populations in Ethiopia: evaluating the role of geography and host specificity in shaping population structure. *International Journal of Plant Sciences* 172: 773-782.

YEH, F.C., R.-C. YANG, T. BOYLE, Z.-H. YE, AND J.X. MAO. 1997. POPGENE: the user-friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Canada. Available at http://www.ualberta.ca/~fyeh/

TABLES AND FIGURES

Table 2.1. Genotyped populations of *Amaranthus tuberculatus*, with name, U.S. state/Canadian province, locality, GPS coordinates, type of population (agricultural or natural), voucher specimen fruit dehiscence (whether the ripe utricle opens or not), and number of individuals collected/genotyped (N).

Name	State/ Province	Locality	Latitude	Longitude	Type	Voucher fruit dehiscence	N
TR	NE	Douglas Co.: Near Waterloo, Two Rivers State Rec Area, on Platte River.	41.223310	96.357610	riverbank	dehiscent	9
CHE	KS	Reno Co.: Cheney State Park, along shore of Cheney Reservoir.	37.744750	97.783860	lake shore	too young to tell	10
TCL	KS	Riley Co.: North of Manhattan, Tuttle Creek Lake State Park.	39.439230	96.710250	lake shore	dehiscent	10
SaltR	OK	Alfalfa Co.: Just outside Salt Plains NWR, over Salt Fork of the Arkansas River.	36.771660	98.038000	riverbank	dehiscent	10
GASC	MO	Gasconade Co.: Gasconade Park along Gasconade River.	38.668122	91.556135	riverbank	too young to tell	10
MSH	MO	Saline Co.: Between Marshall and Malta Bend, just off Hwy 65 in soy field.	39.168783	93.289057	soy field	dehiscent	10
AAF	MO	St. Louis Co.: Eureka, across from Allenton Access to Meramec River in old flooded field.	38.473250	90.661016	soy field	dehiscent	10
WSR	MO	St. Charles Co.: Defiance, Weldon Spring Conservation Area off Hwy 94, banks of Missouri River.	38.656280	90.736950	riverbank	indehiscent	10
GTP	MO	St. Louis Co.: Kirkwood, Green Tree Park on Marshall Road, banks of Meramec River.	38.558930	90.447360	riverbank	dehiscent	10
WSS	MO	St. Charles Co.: Defiance, Weldon Spring Conservation Area off Hwy 94, sunflower fields near Missouri River.	38.656280	90.736950	sunflower field	dehiscent	7
EMN	MO	St. Louis Co.: Kirkwood, Emenegger Nature Park, along Meramec River.	38.545160	90.433450	riverbank	dehiscent	8
JCK	AR	Jackson Co.: Near Newport, Jacksonport State Park, on bank of White River.	35.642113	91.319192	riverbank	dehiscent	9
PEK	IL	Peoria Co.: Pekin, left bank of the Illinois River under the Hwy 9 bridge.	40.574410	89.655980	riverbank	indehiscent	10
RIP	IL	Schuyler Co.: Ripley, on La Moine River.	40.027434	90.631546	riverbank	indehiscent	10
KEY	IL	Clinton Co.: Keyesport Recreation Area, Carlyle Reservoir.	38.733710	89.275850	lake shore	too young to tell	10
KEYC	IL	Clinton Co.: North of Keyesport along Mulberry Rd, in old field.	38.768113	89.273209	soy field	dehiscent	10
KNK	IL	Kankakee Co.: 2 miles east of Momence, soybean field.	41.160983	87.627515	soy field	indehiscent	10
VIGO	IN	Vigo Co.: margin of soybean field along IN 246.	39.273930	87.470000	soy field	indehiscent	10
KANK	IN	Starke Co.: Kankakee FWR, off of Hwy 8 and 39, Kankakee River.	41.314810	86.737550	riverbank	indehiscent	10
WAB	IN	Wabash Co.: Wabash, banks of the Wabash River.	40.790980	85.820860	riverbank	indehiscent	10
IND	IN	Marion Co.: Indianapolis, Left Fork of the White River, on bank.	39.783310	86.189750	riverbank	indehiscent	10
AUR	IN	Dearborn Co.: Aurora, boat ramp E of Hwy 56 on Ohio River.	39.056110	84.898350	riverbank	too young to tell	10
BTL	OH	Butler Co.: Hamilton, Veteran's Field Park, on Great Miami River.	39.427430	84.540710	riverbank	dehiscent	9
PCL	OH	Highland Co.: Paint Creek Lake State Park, past dam, on lake shore.	39.268010	83.388610	lake shore	indehiscent	10
NEV	OH	Clermont Co.: Neville, boat ramp and bank along the Ohio River.	38.807630	84.211710	riverbank	indehiscent	10
STW	OH	Miami Co.: Covington City Park, on the bank of the Stillwater River.	40.121630	84.358660	riverbank	indehiscent	10
CAN	OH	Madison Co.: South of Plain City along Hwy 42, soybean field.	39.985850	83.339630	soy field	indehiscent	10
SCIO	OH	Delaware Co. : O'Shaughnessy Reservoir along Scioto River.	40.177450	83.126400	lake shore	dehiscent	10
GTB	OH	Miami Co.: East of Gettysburg on Hwy 36, in soybean field.	40.120100	84.398680	soy field	indehiscent	10

Name	State/ Province	Locality	Latitude	Longitude	Type	Voucher fruit dehiscence	N
MC	OH	Union Co.: Milford Center, soybean field south along Hwy 36.	40.155580	83.455330	soy field	indehiscent	10
RT29	OH	Mercer Co.: West of Celina, Rt. 29 cornfield.	40.545911	84.634131	corn field	dehiscent	10
OTT	OH	Putnam Co.: Between Ottawa and Findlay, off of Hwy 224 on the Blanchard River.	41.037830	83.813490	riverbank	indehiscent	10
MAU	OH	Lucas Co.: Maumee, Side Cut Metropark along Maumee River.	41.556350	83.662410	riverbank	indehiscent	10
PTC	OH	Ottawa Co.: Port Clinton, along beach in Municipal Pier area, on Lake Erie.	41.514500	82.938430	lake shore	indehiscent	9
DMD	MI	Eaton Co.: Dimondale, on bank of Grand River near bridge across Bridge St.	42.645000	84.649700	riverbank	indehiscent	9
DEL	ON	Middlesex Co.: Near Delaware, Thomas River on Co. Rd. 16.	42.933750	81.421060	riverbank	indehiscent	9
SCF	ON	Essex Co.: Near Leamington, Seacliffe Park, along beach.	42.030950	82.603850	lake shore	too young to tell	10
YORK	ON	Haldimand Co.: York, south of Caledonia. Along bank of Grand River.	43.020700	79.891050	riverbank	indehiscent	10

Table 2.2. Microsatellite loci forward (F) and reverse (R) primers, repeat motif, dye label (used in multiplexed reactions), size range, and primer source. Sources are Lee et al. (2009) and the Tranel lab at the University of Illinois-Urbana/Champaign.

Locus name	Primer	Repeat	Dye label	Size range	Source
C1140	F: 5'-TTGAAGACGACGATCTTTCTGGAT	$(GAT)_{10}$	6FAM	113-181 bp	Tranel lab *A. tuberculatus* transcriptome data
	R: 5'-CCCCTCTGTACACCATAATCGAAC				
C4097	F: 5'-ATCATCTTCTGCTAAGGCTGTTGG	$(ACC)_8$	NED	164-179 bp	Tranel lab *A. tuberculatus* transcriptome data
	R: 5'-ATATCTTCCCCAATTGGACTCCTC				
C0745	F: 5'-TAGGAAGTTCATCCATAAGCTCGG	$(TGA)_{10}$	NED	130-164 bp	Tranel lab *A. tuberculatus* transcriptome data
	R: 5'-CAATTCCAAGGAATCATCCTCATC				
C3561	F: 5'-CCATAAACCATTTTCCCAGACC	$(CCA)_8$	HEX	123-141 bp	Tranel lab *A. tuberculatus* transcriptome data
	R: 5'-ACTTCTGGCCCAATTAGGAAGTC				
C4999	F: 5'-CCACCCAATGACCCATACCTACTA	$(ACC)_8$	NED	120-141 bp	Tranel lab *A. tuberculatus* transcriptome data
	R: 5'-GATGAGGTTGATAATTGGGGTTCA				
AAC1	F: 5'-CCCACCAAGGATGATCATTTAGAC	AAC	6FAM	112-130 bp	Lee et al. 2009
	R: 5'-TCATCATTATTTGTTGGCGTTGAC				
TAG5	F: 5'-GTCGCTGAATTGTTTTAGCTTGGT	TAG	HEX	132-163 bp	Lee et al. 2009
	R: 5'-TGGGAATTCTCTCTTGTGACACAGT				
ATC9	F: 5'-TAGCCATTTCAACCTTACGAGGAA	ATC	NED	142-160 bp	Lee et al. 2009
	R: 5'-ACCGTTGATTGATTTTATGGCATC				
C3695	F: 5'-TCAACTTCTTATTCTTGGGTTGCTTC	$(TGA)_8$	6FAM	127-174 bp	Tranel lab *A. tuberculatus* transcriptome data
	R: 5'-CCTTACCTTCTCTCAAAAGCACCA				
C9333	F: 5'-AACTAAACGCATTTGCCATTGAA	$(GAT)_8$	HEX	165-199 bp	Tranel lab *A. tuberculatus* transcriptome data
	R: 5'-TGTTCATCTAACCACATCATAATGGAA				

Table 2.3. Population genetic statistics for each population summarized over all loci. Na = number of alleles, Ne = effective number of alleles (estimated reciprocal of homozygosity), Ho = observed heterozygosity, He = expected heterozygosity.

Population	Mean Na	Mean Ne	Mean Ho	Mean He
TR	5.1	3.22	0.5722	0.6537
CHE	4.9	3.48	0.5679	0.5971
TCL	5.2	3.05	0.5856	0.5937
SaltR	5.2	3.50	0.6544	0.6829
GASC	4.9	3.09	0.5425	0.5428
MSH	4.8	3.07	0.5044	0.5497
AAF	4.7	3.39	0.6050	0.6069
WSR	4.9	3.39	0.5667	0.6121
GTP	5.1	3.52	0.5933	0.6211
WSS	4.4	3.44	0.6271	0.6430
EMN	4.6	3.21	0.4857	0.5463
JCK	5.2	3.70	0.5139	0.6165
PEK	4.8	3.31	0.5956	0.5777
RIP	5.1	3.12	0.5400	0.5663
KEY	5.0	3.27	0.5500	0.6000
KEYC	4.7	3.17	0.5100	0.5642
KNK	5.3	3.18	0.5522	0.5955
VIGO	4.3	2.56	0.6900	0.5531
KANK	4.8	3.22	0.6100	0.6016
WAB	5.3	3.44	0.6300	0.6396
IND	5.6	4.05	0.5956	0.6624
AUR	4.4	3.01	0.4611	0.5379
BTL	4.7	3.00	0.5111	0.5850
PCL	3.9	2.86	0.4878	0.5054
NEV	3.8	2.64	0.5244	0.4895
STW	5.0	3.55	0.5667	0.5974
CAN	3.3	2.34	0.5819	0.5531
SCIO	4.5	3.27	0.6400	0.6253
GTB	4.5	2.74	0.4733	0.5290
MC	3.7	2.44	0.6197	0.5781
RT29	5.0	3.64	0.6000	0.6393
OTT	4.2	2.83	0.4873	0.5862
MAU	4.5	3.08	0.5444	0.5980
PTC	4.6	2.96	0.5861	0.5851
DMD	4.2	2.91	0.4986	0.5521
DEL	3.0	2.01	0.4560	0.4599
SCF	4.7	3.47	0.6033	0.6282
YORK	3.5	2.01	0.3818	0.4245
All Populations	4.6	3.11	0.5557	0.5816

Table 2.4. Population genetic statistics for each locus summarized over all populations. Na = number of alleles, Ne = effective number of alleles (estimated reciprocal of homozygosity), Ho = observed heterozygosity, He = expected heterozygosity.

Locus	Sample size	Na	Ne	Ho	He
C1140	722	20	8.86	0.8227	0.8884
C4097	720	6	2.04	0.4861	0.5094
C0745	702	14	5.87	0.7578	0.8309
C3561	722	7	1.55	0.3518	0.3535
C4999	716	9	3.02	0.5782	0.6695
AAC1	714	6	2.46	0.4342	0.5941
TAG5	710	10	2.55	0.4620	0.6086
ATC9	716	6	1.09	0.0894	0.0867
C3695	718	16	9.10	0.8468	0.8913
C9333	704	13	6.46	0.7273	0.8464

Table 2.5. Results of the AMOVA analyses of the 38-population total species range dataset, the 12-population Ohio dataset, and the 5-population St. Louis dataset. The AMOVA regions are agricultural populations vs. non-agricultural populations in the "crop vs. non-crop" analyses, and populations assigned by STRUCTURE primarily to one genetic cluster or the other for the "STRUCTURE-based genetic regions" analyses. Df = degrees of freedom. "P-value" comes from 999 permutations to estimate the range of values for a dataset with no subdivision.

Total Species Range Dataset								
Crop vs. Non-Crop Regions					**STRUCTURE-based Genetic Regions**			
Source	df	Variance	%		Source	df	Variance	%
Among Regions	1	0.002	0%		Among Regions	1	0.167	5%
Among Pops	36	0.228	7%		Among Pops	36	0.164	5%
Within Pops	700	3.021	93%		Within Pops	700	3.003	90%
Statistic	Value	P-value			Statistic	Value	P-value	
Frt	0.000	0.232			Frt	0.050	0.001	
Fsr	0.070	0.001			Fsr	0.052	0.001	
Fst	0.071	0.001			Fst	0.099	0.001	

Ohio Dataset								
Crop vs. Non-Crop Regions					**STRUCTURE-based Genetic Regions**			
Source	df	Variance	%		Source	df	Variance	%
Among Regions	1	0.042	1%		Among Regions	1	0.082	3%
Among Pops	10	0.197	6%		Among Pops	10	0.172	5%
Within Pops	224	2.965	93%		Within Pops	224	2.965	92%
Stat	Value	P-value			Statistic	Value	P-value	
Frt	0.013	0.001			Frt	0.025	0.001	
Fsr	0.062	0.001			Fsr	0.055	0.001	
Fst	0.074	0.001			Fst	0.079	0.001	

St. Louis Dataset			
Crop vs. Non-Crop Regions			
Source	df	Variance	%
Among Regions	1	0.000	0%
Among Pops	3	0.118	4%
Within Pops	85	3.087	96%
Statistic	Value	P-value	
Frt	-0.004	0.727	
Fsr	0.037	0.001	
Fst	0.033	0.001	

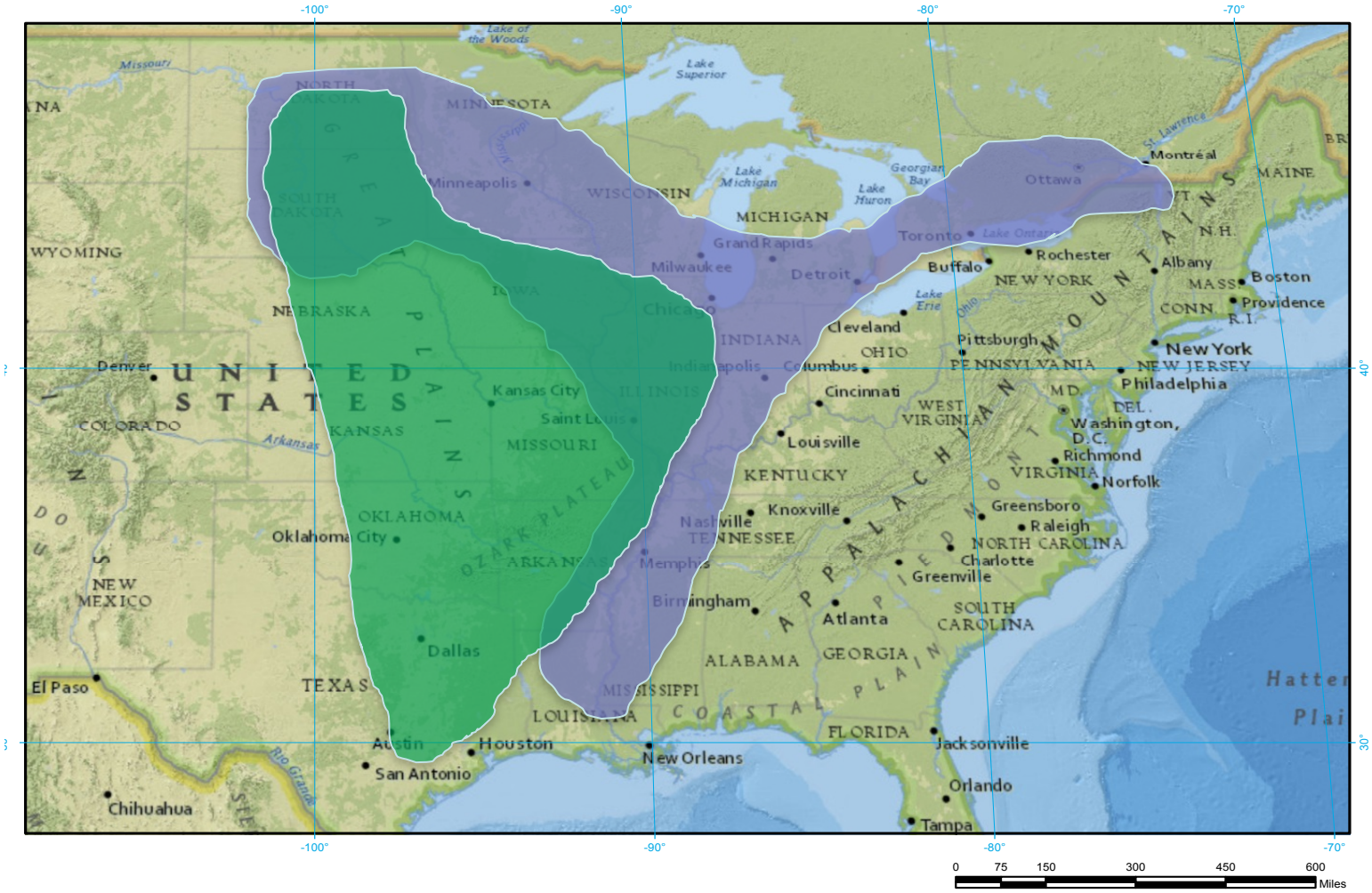

Figure 2.1. Geographical range of *Amaranthus tuberculatus* s.l. (waterhemp), with historical range of *A. tuberculatus* var. *rudis* in green, and range of *A. tuberculatus* var. *tuberculatus* in purple, with the opaque green shading showing the areas of overlap between the varieties (adapted from Sauer, 1957). The map is the National Geographic Basemap in ArcGIS.

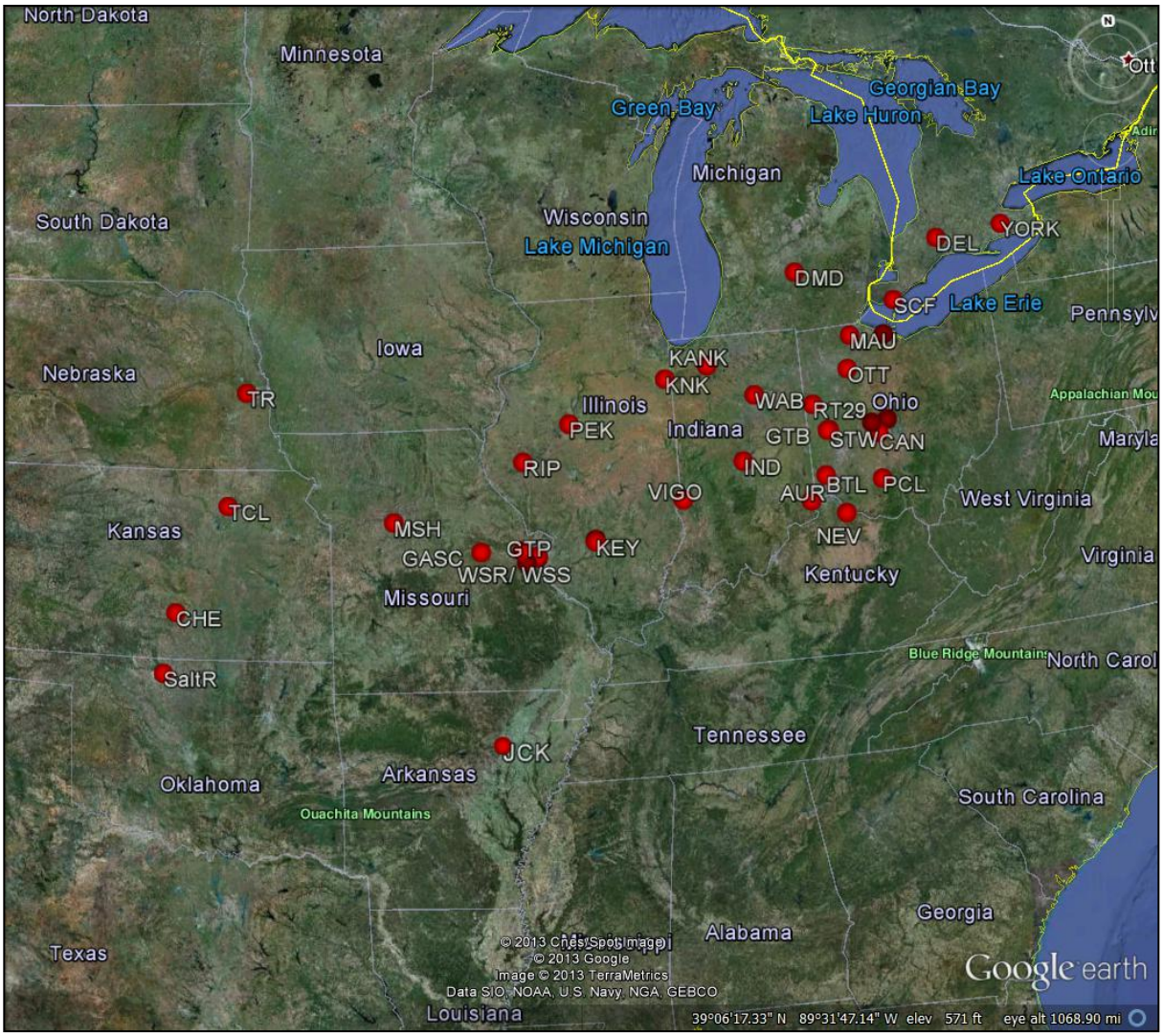

Figure 2.2. Locations of 38 genotyped populations of *Amaranthus tuberculatus* from across the entire species range. Geographic coordinates were plotted in Google Earth.

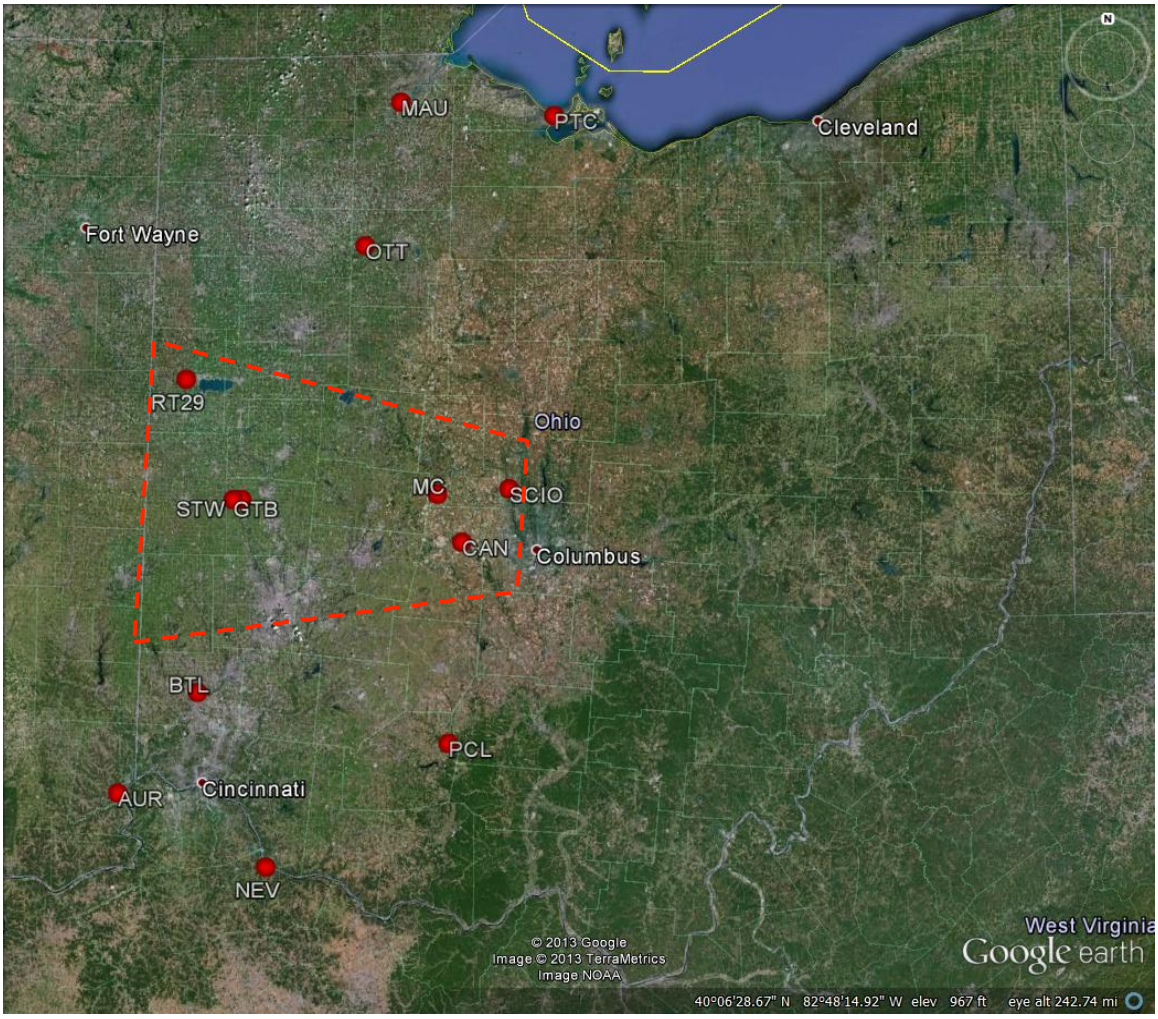

Figure 2.3. Locations of 12 genotyped populations of *Amaranthus tuberculatus* from Ohio. Geographic coordinates were plotted in Google Earth. The red dashed trapezoid outlines the "agricultural waterhemp region" of the state.

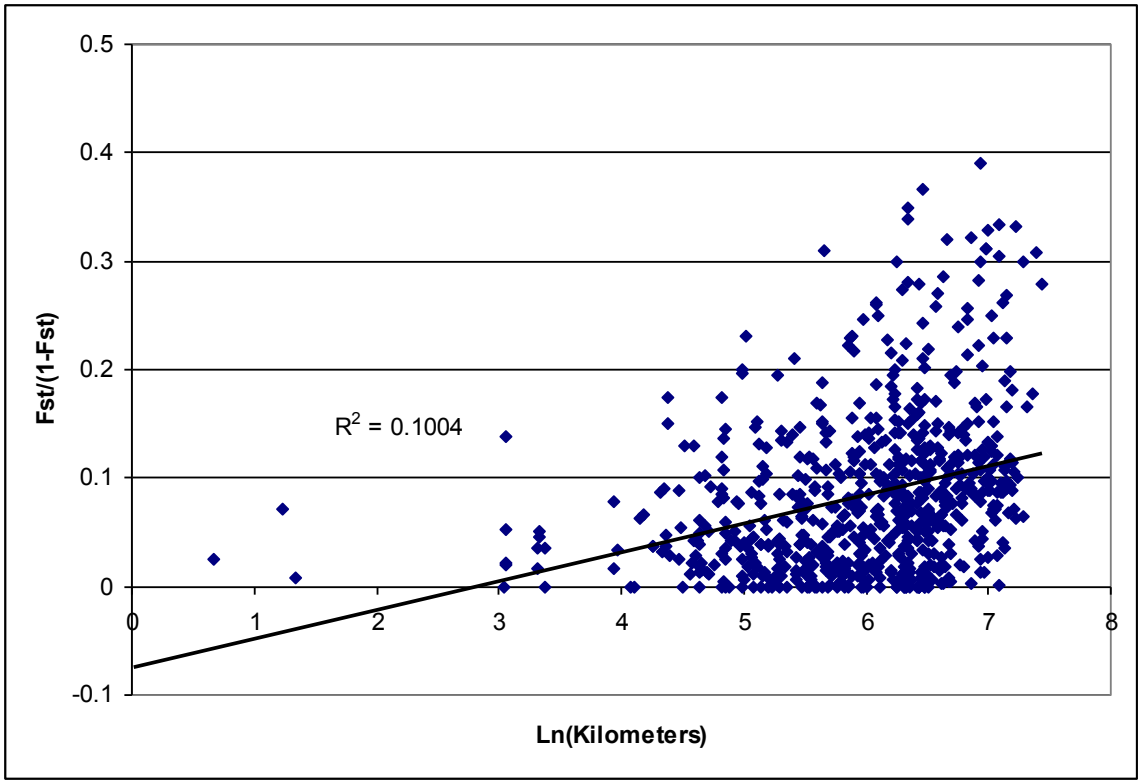

Figure 2.4. Plot of pairwise genetic distances (Fst/(1-Fst)) versus pairwise geographic distances (ln(kilometers)) for the 38 *Amaranthus tuberculatus* populations genotyped over the entire species range, showing isolation by distance.

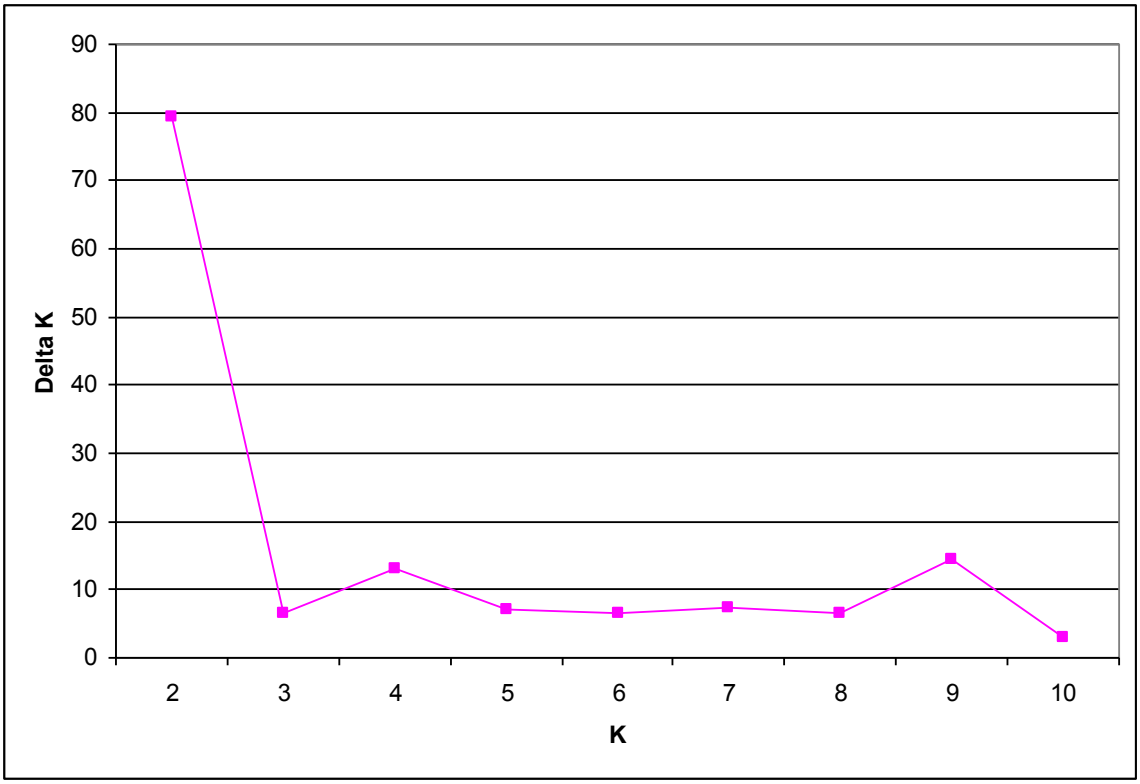

Figure 2.5. Plot of Delta K (the second-order rate of change of K) versus K, for the 38 genotyped populations from across the species range, showing a highest delta K of 2. Delta K is calculated from ln Prob(X|K) values for each of three runs at K=n in STRUCTURE, following Evanno et al., 2005.

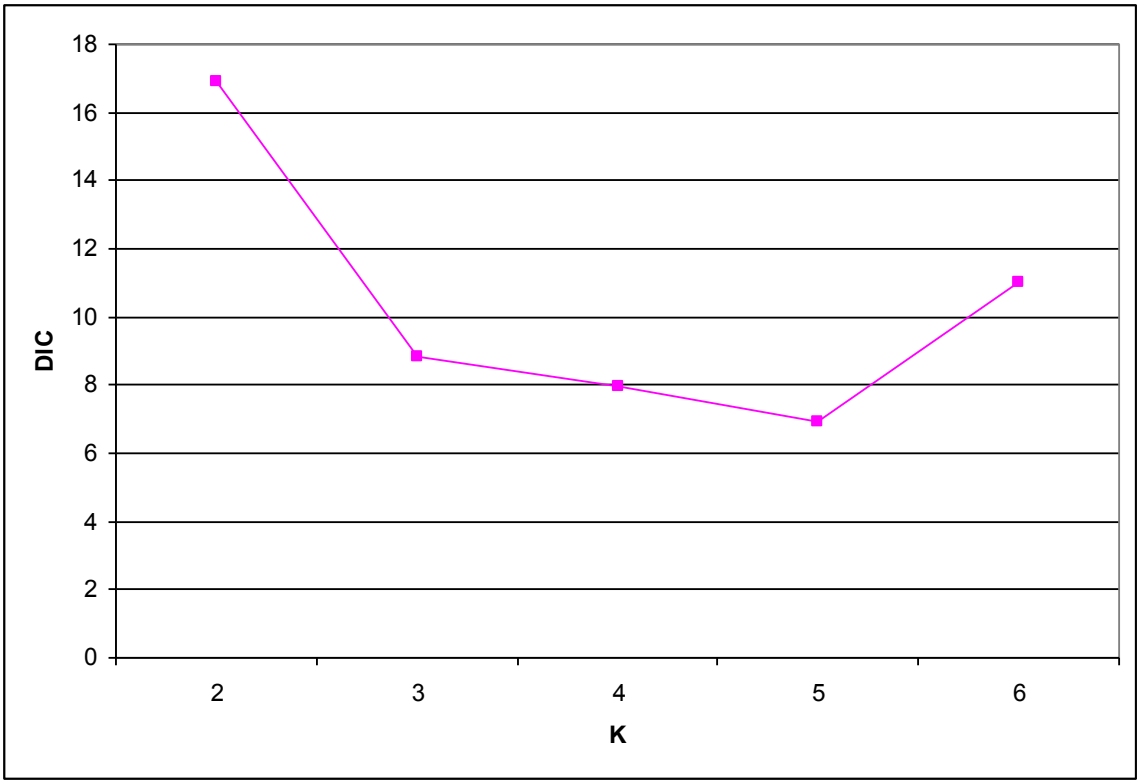

Figure 2.6. Plot of Delta K (the second-order rate of change of K) versus K, for the 12 genotyped populations from Ohio, showing a highest delta K of 2. Delta K is calculated from ln Prob(X|K) values for each of three runs at K=n in STRUCTURE, following Evanno et al., 2005.

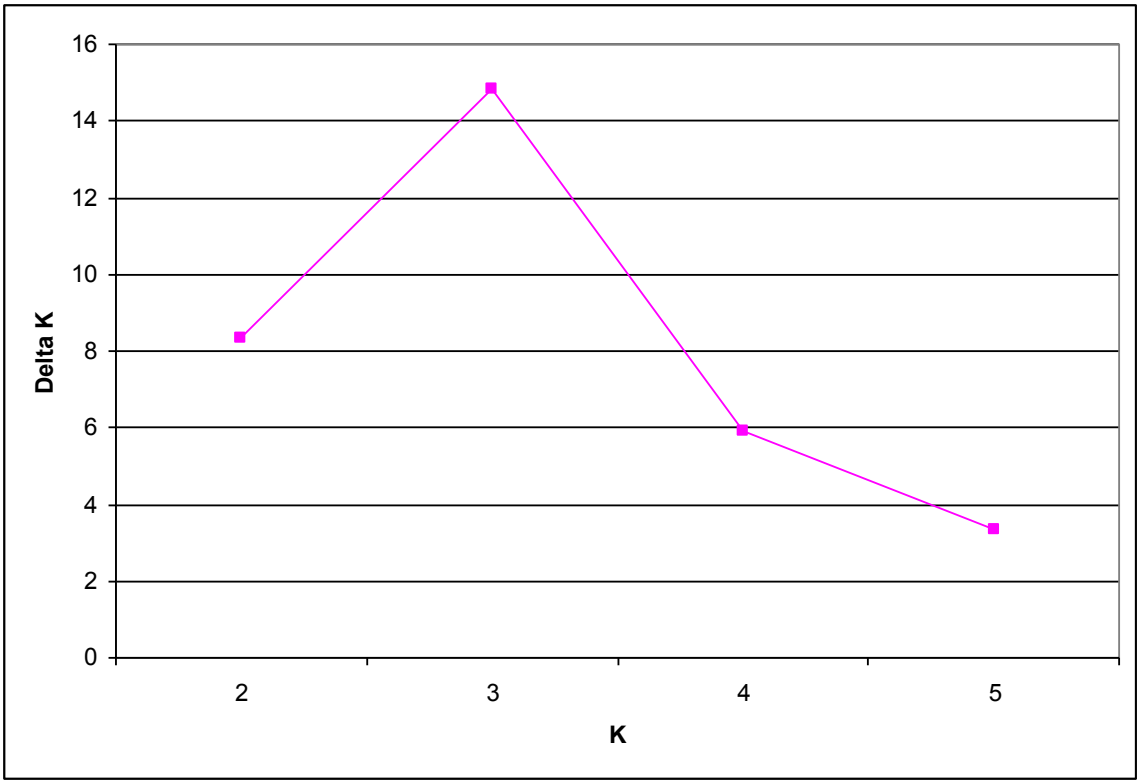

Figure 2.7. Plot of Delta K (the second-order rate of change of K) versus K, for the 5 genotyped populations from St. Louis. Delta K is calculated from ln Prob(X|K) values for each of three runs at K=n in STRUCTURE, following Evanno et al., 2005. K is probably actually equal to 1, as shown below in Figure 2.11.

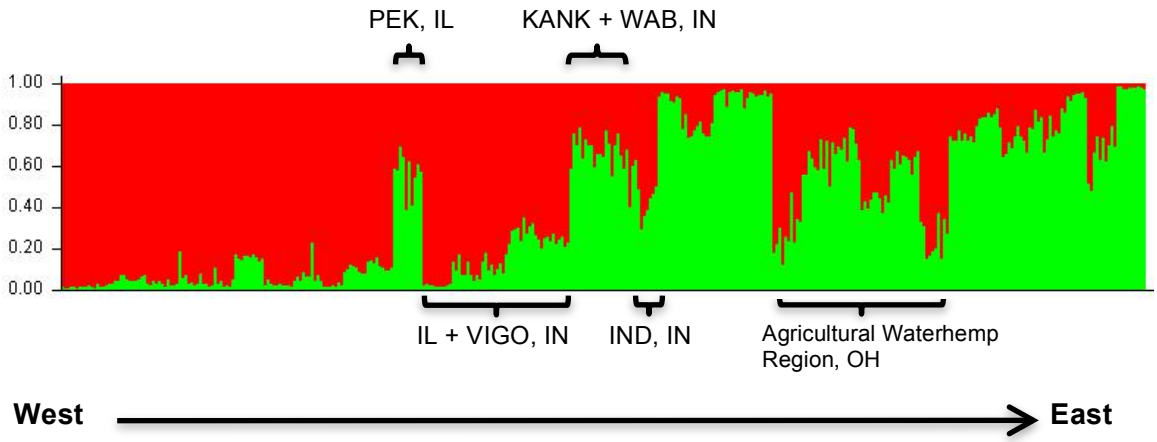

Figure 2.8. STRUCTURE bar graph for K=2, for the 38 genotyped populations from across the species range, showing assignment of individuals (vertical lines) to two genetic clusters (shown by the colors). The colored segments of each individual show the proportion of its assignment to each genetic cluster. The "western" genetic cluster is in red and the "eastern" genetic cluster is in green. Populations of interest are shown with brackets and names above and below the bar graph, and the organization of the populations geographically is shown by the arrow below the graph.

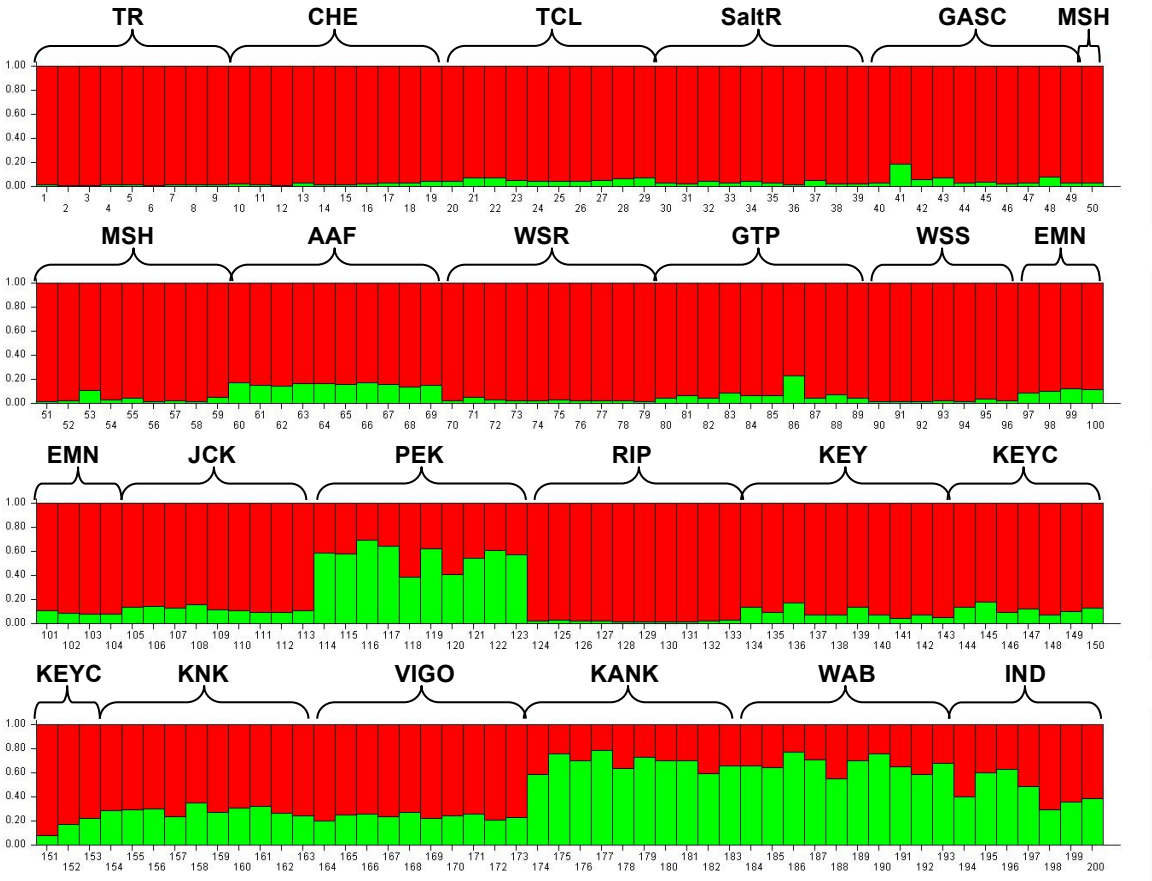

Figure 2.9a. Closer view of the first half of a STRUCTURE bar graph for K=2, for the 38 genotyped populations from across the species range. The "western" genetic cluster is in red and the "eastern" genetic cluster is in green. Populations are shown with brackets and names above the bar graph.

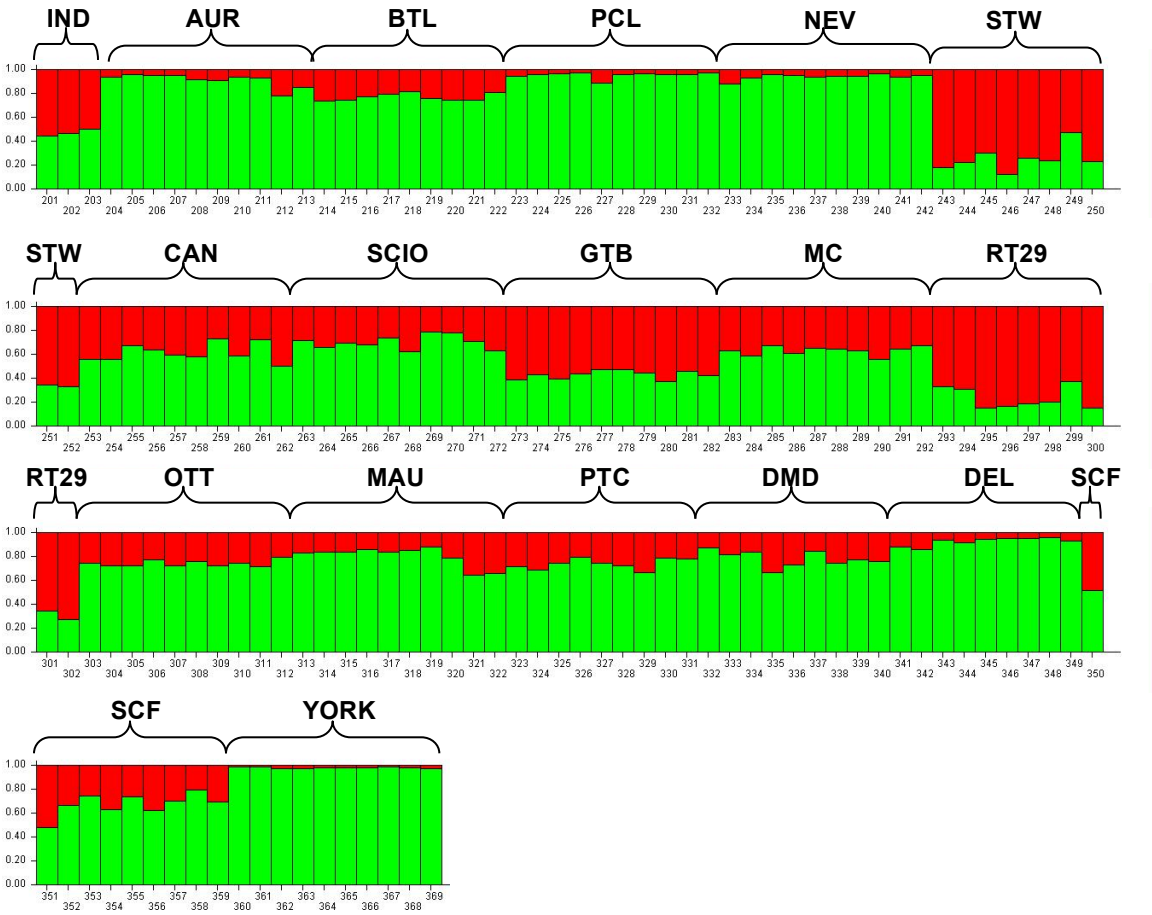

Figure 2.9b. Closer view of the second half of a STRUCTURE bar graph for K=2, for the 38 genotyped populations from across the species range. The "western" genetic cluster is in red and the "eastern" genetic cluster is in green. Populations are shown with brackets and names above the bar graph.

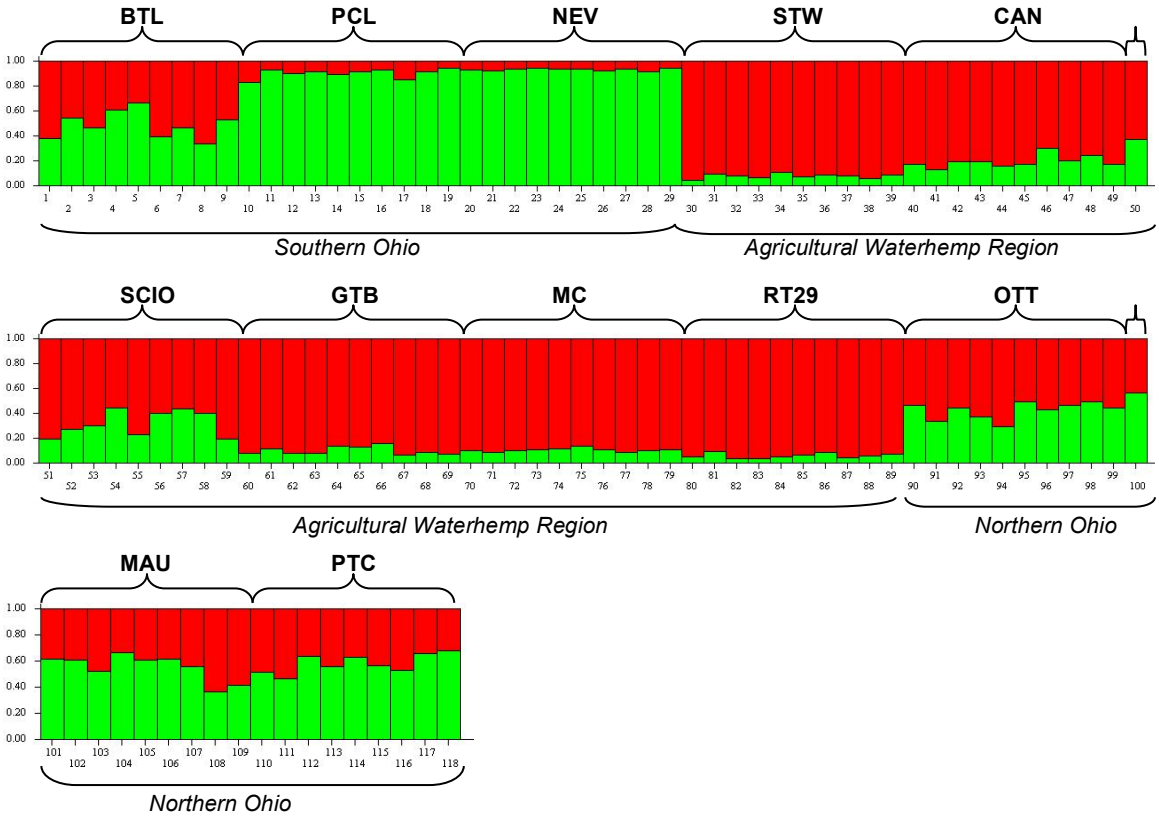

Figure 2.10. STRUCTURE bar graph for K=2, for the 12 genotyped populations from Ohio. The "western" genetic cluster is in red and the "eastern" genetic cluster is in green. Populations are shown with brackets and names above the bar graph, and geographical regions of interest are shown with brackets and names below the bar graph.

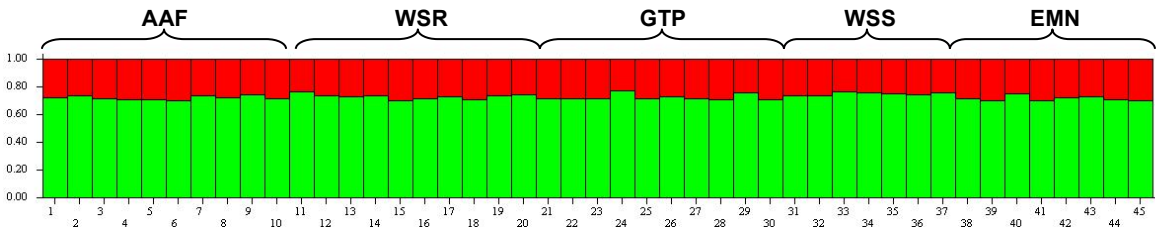

Figure 2.11. A representative STRUCTURE bar graph for K=2 for the 5 genotyped populations from St. Louis. Note the nearly equal subdivision of every individual between the two genetic clusters, indicating no population substructure and a real K of 1. Populations are shown with brackets and names above the bar graph.

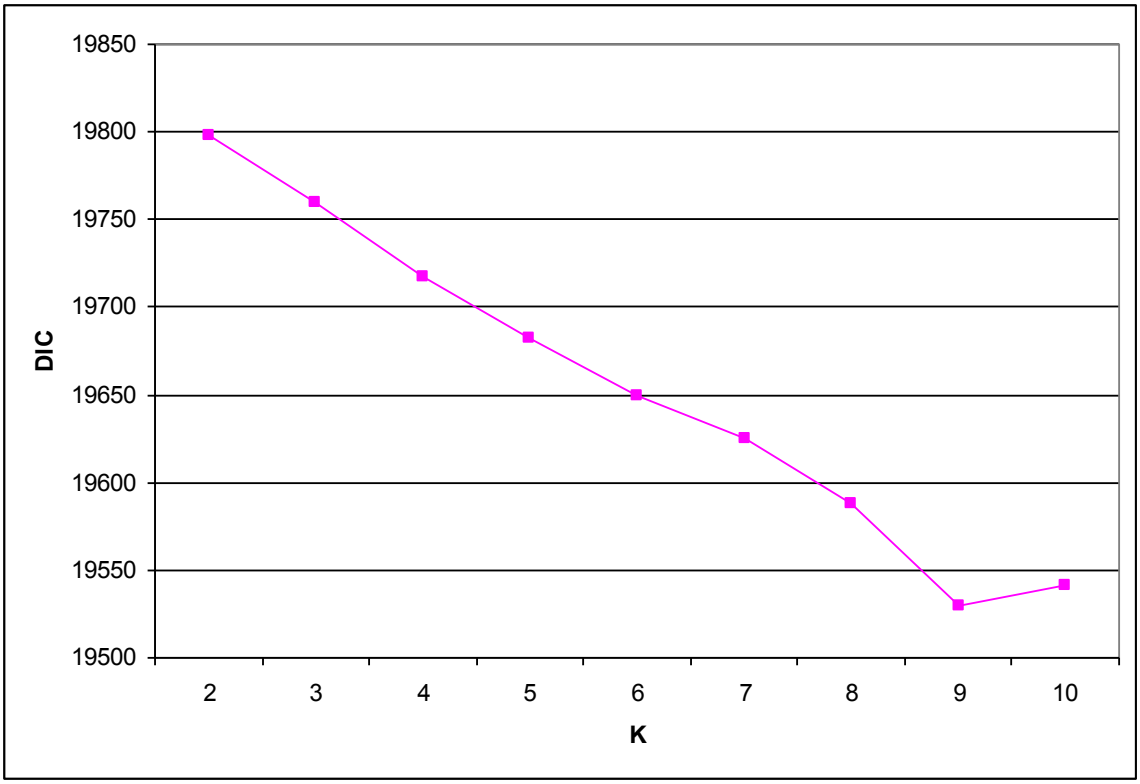

Figure 2.12. Plot of DIC (the deviance information criterion) versus K, for the 38 genotyped populations from across the species range. DIC was calculated by the program TESS and averaged over 3 runs at each value of K. The DIC values are not as informative for this dataset as the bar graphs and Voroni tessellation diagrams, shown in Figures 2.13 and 2.14.

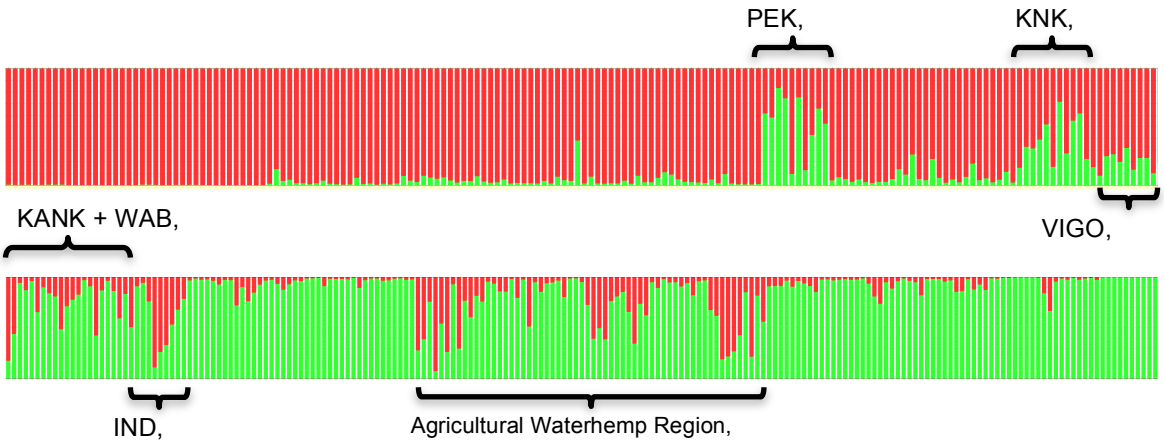

Figure 2.13. TESS bar graph for K=2, for the 38 genotyped populations from across the species range, showing assignment of individuals (vertical lines) to two genetic clusters (shown by the colors). The colored segments of each individual show the proportion of its assignment to each genetic cluster. The "western" genetic cluster is in red and the "eastern" genetic cluster is in green. Populations of interest are shown with brackets and names above and below the bar graph, and the organization of populations geographically is the same as in Figure 2.8.

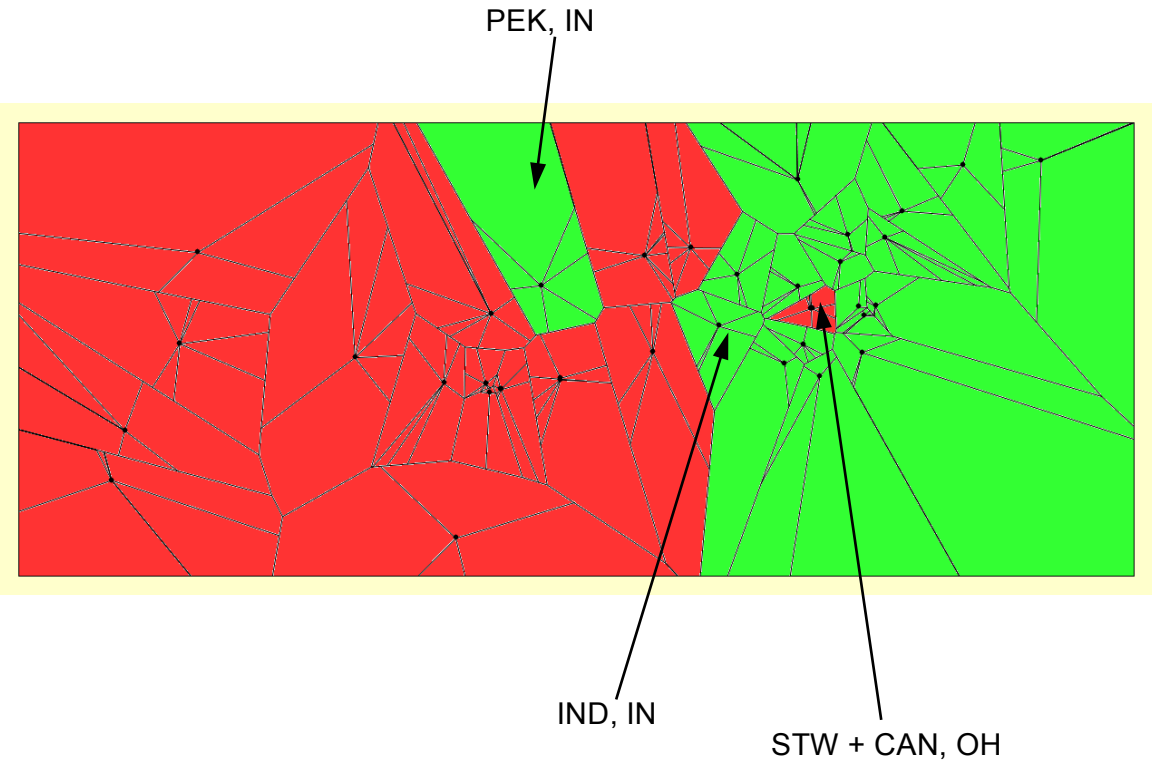

Figure 2.14. TESS Voroni tessellation “hard clustering” diagram for K=2, for the 38 genotyped populations from across the species range, showing the genetic clusters with different colors. The “western” genetic cluster is in red and the “eastern” genetic cluster is in green. Populations of interest are labeled.

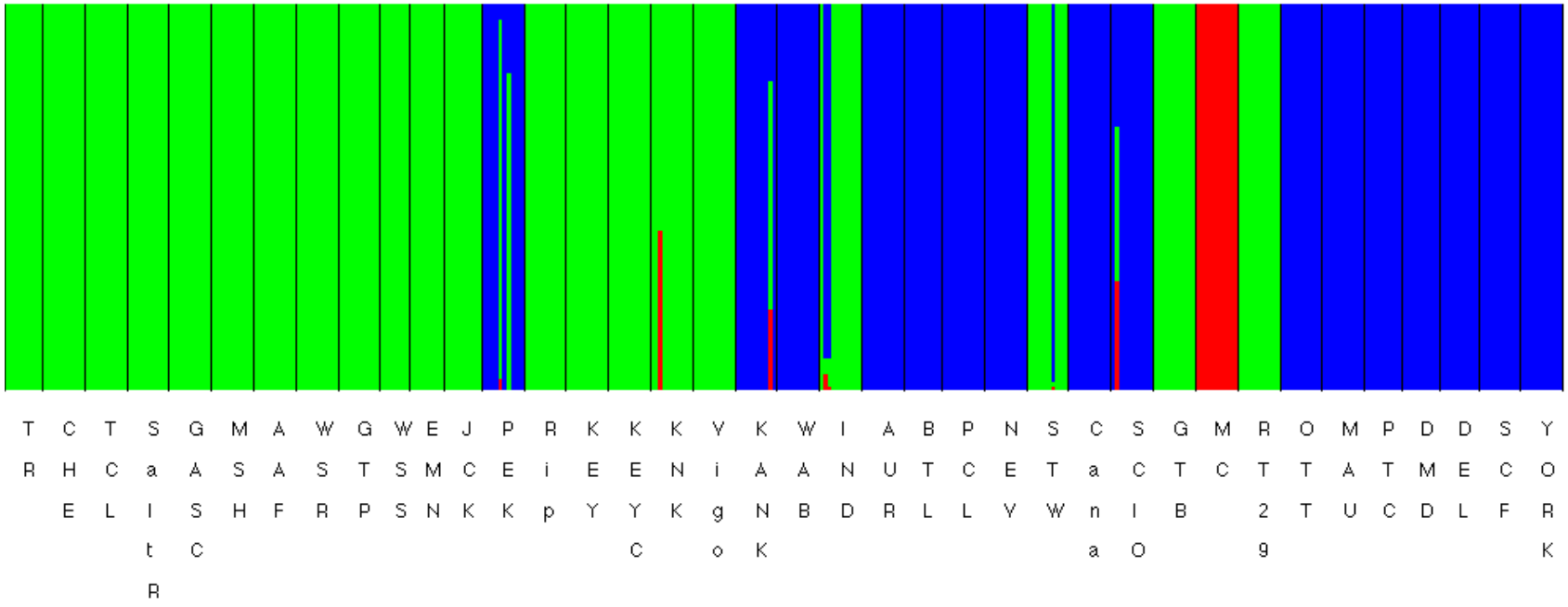

Figure 2.15. BAPS admixture bar graph for K=3, for the 38 genotyped populations from across the species range, showing assignment of individuals (vertical lines) to three genetic clusters (shown by the colors). The colored segments of each individual show the proportion of its assignment to each genetic cluster. The "western" genetic cluster is in green and the "eastern" genetic cluster is in blue, while the third cluster (the MC population) is in red. Population names are shown below the population clusters, and the organization of populations geographically is the same as in Figures 2.8 and 2.13.

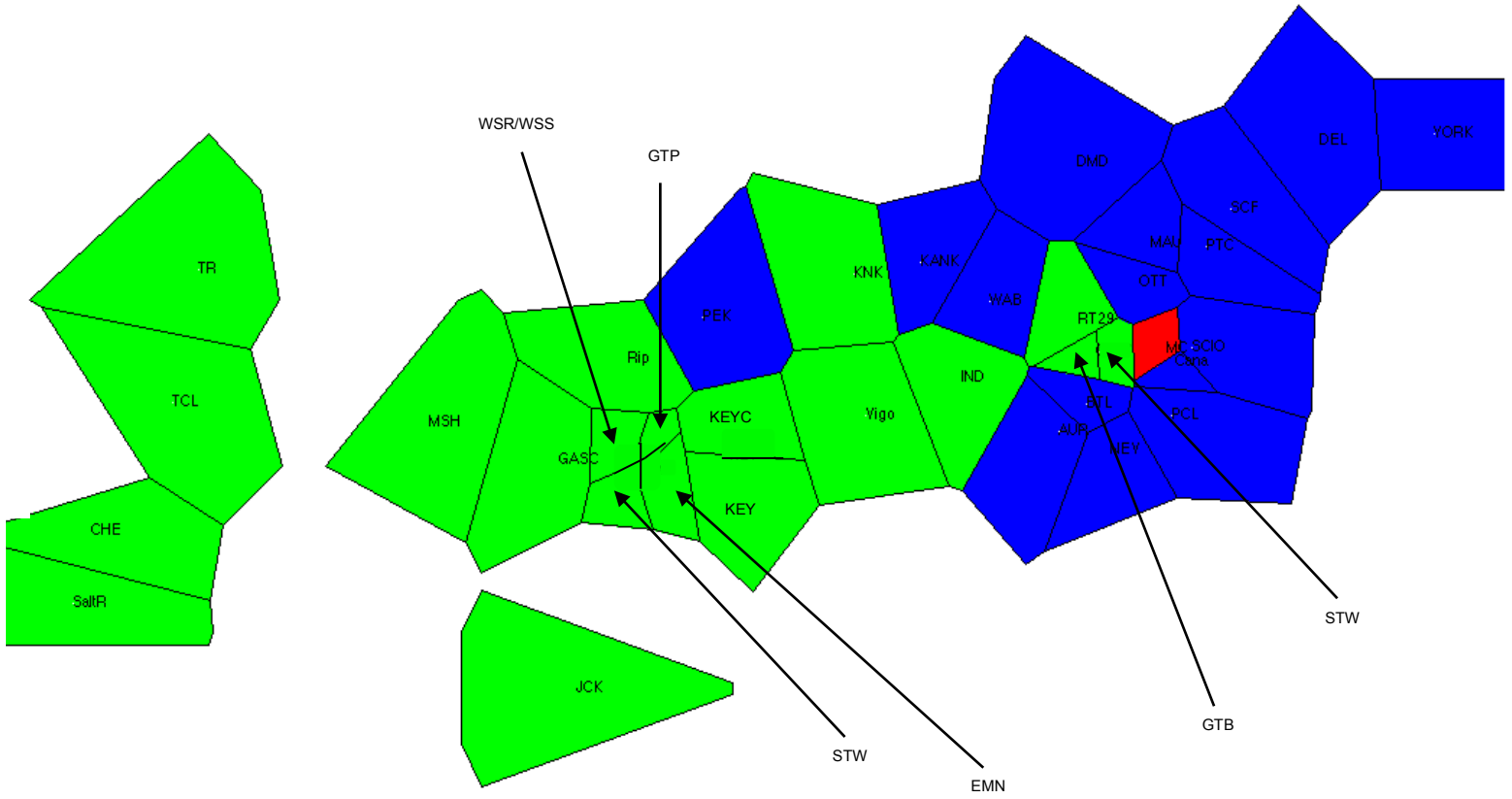

Figure 2.16. BAPS Voroni tessellation diagram for K=3, for the 38 genotyped populations from across the species range, showing the genetic clusters with different colors. The “western” genetic cluster is in green and the “eastern” genetic cluster is in blue, while the third cluster (the MC population) is in red. Populations are labeled.

CHAPTER 3

Agricultural and Local Adaptation Shape Invasiveness in Waterhemp, *Amaranthus tuberculatus* (Amaranthaceae)

INTRODUCTION

The importance of adaptive evolution in the spread of invasive species has only recently been recognized (Sakai et al., 2001). In the last decade, experiments and models have begun to demonstrate that invasive species of natural habitats can exhibit adaptive evolution in their new range, associated with establishment and range expansion (reviewed by Lee, 2002; Lambrinos, 2004). Local adaptation has been included in models of invasion speed (García-Ramos and Rodríguez, 2002), and documented empirically in some invasive species (e.g., Huey et al., 2000; Maron et al., 2004; Scott et al., 2010). A number of studies have also observed increased genetic variation or evolutionary potential from population admixture after multiple introductions of an invasive species (e.g., Novak and Mack, 1993; Kolbe et al., 2004; Lavergne and Molofsky, 2007), or interspecific hybridization (e.g., Gaskin and Schaal, 2002; Ayres et al., 2004; reviewed in Dlugosch and Parker, 2008; Schierenbeck and Ellstrand, 2009). However, several counter-examples have been found, where invasive species have persisted and expanded in their introduced range despite a lack of genetic variation and/or local adaptation (e.g., Parker et al., 2003; Loomis and Fishman, 2009; Ebeling et al., 2011).

The role of local adaptation in agricultural weed invasions is particularly poorly understood (Clements et al., 2004; Neve et al., 2009; Vigueira et al., 2012). Agricultural weeds are invasive plants of heavily human-modified environments, namely crop fields and rangelands. With the exception of herbicide resistance, where selection for resistance mutations is extensively documented (e.g. Davis et al., 2009; Tranel and Wright, 2009; Délye et al., 2010; reviewed by Owen and Zelaya, 2005; Powles and Yu, 2010), the mechanisms permitting agricultural weed invasions, and the role of adaptive evolution in this process, are almost completely unknown. The potential for weed evolution in response to agricultural selection

pressures has been recognized since at least 1965, when Harlan and De Wet proposed that agricultural weeds were as heavily influenced by anthropogenic selection as domesticated species. In a later paper, these authors pointed out that crop field weeds can contain intraspecific variation in adaptation to agricultural environments (De Wet and Harlan, 1975). Their ideas about weeds differed from those of their contemporary Baker, whose 1974 treatise on "The Evolution of Weeds" emphasized the concept of a "general-purpose genotype," which envisioned weed species as strongly phenotypically plastic, thereby negating the need for local adaptation to different abiotic conditions. Since the 1970s, experiments have shown that many agricultural weeds do contain adaptive genetic variation and may respond to selection (reviewed by Vigueira et al., 2013). Agricultural weeds cost an estimated $33 billion annually in the U.S. alone (Pimentel et al., 2005); given their economic importance, it is surprising that adaptive evolution of weeds has not been more frequently studied.

Like invasive species, agricultural weeds may contain high levels of genetic variation due to multiple introductions or origins, and they may be locally adapted to different environmental (agricultural) conditions (e.g., Lai et al., 2008; Muller et al., 2009; Délye et al., 2010; Reagon et al., 2010). Unlike most invasive species, however, an agricultural weed may be native to the geographical region where it is invasive: many weed species were originally pioneers in naturally disturbed habitats, before making the ecological and/or evolutionary leap to crop fields (De Wet and Harlan, 1965). The fact that both native and introduced weeds often occur outside agricultural fields, in natural environments or as ruderal weeds of railroads and roadsides, adds to the opportunities for local adaptation in these species.

Research on local adaptation of invasive species often includes common garden or reciprocal transplant experiments. These experiments may be designed to examine adaptive

differences between populations in the native and introduced range of a species (Williams et al., 2008; Eriksen et al., 2012), local adaptation to different environments in the introduced range (Leger and Rice, 2007), or both (Hodgins et al., 2011, 2013). Despite the potential of weeds as evolutionary systems, common garden and reciprocal transplant experiments have rarely been used to study adaptation in agricultural weeds (but see Keller et al., 2000; Leiss and Müller-Scharer, 2001; Bommarco et al., 2010).

My study system is the Midwestern species *Amaranthus tuberculatus* (Amaranthaceae), commonly called waterhemp (see Chapter 2 for study system details). Since waterhemp is a Midwestern native, which has only recently invaded agricultural ecosystems and has demonstrably evolved in response to changing agricultural practices (i.e., herbicide application), it is an ideal candidate for the study of intraspecific variation in adaptation in agricultural weeds. While *A. tuberculatus* naturally ranges from the Great Plains to southern Ontario, the region of agriculturally-problematic waterhemp is smaller: it is a major cause of crop yield loss in parts of Iowa, Missouri, Illinois, and Indiana, and a lesser problem in the Plains states, Texas, Kentucky, and Ohio (see Figure 3.1) (Tranel and Trucco, 2009). Taxonomists have long observed geographically-structured morphological variation within the species. One researcher distinguished two species within *A. tuberculatus* (Sauer, 1955), which are sometimes still recognized as varieties (Costea and Tardif, 2003). Waterhemp appears to have been diverging into two species on either side of the Mississippi River, which were brought back into contact by the spread of large-scale Midwestern agriculture and the subsequent waterhemp invasion in the 20th century (Sauer, 1957). A population genetic study has shown that, of these varieties, agriculturally invasive waterhemp is strongly associated genetically with the "western" variety

(see Chapter 2). However, it is unclear whether this variety is also phenotypically distinct in ways that make it more adapted to agricultural environments, particularly in the Midwest.

While some previous studies have examined herbicide resistance and other fitness components of waterhemp within agricultural fields with the ultimate goal of weed control (e.g., Hager et al., 2002a; Steckel et al., 2002; Hartzler et al., 2004), no previous study has explicitly compared the relative fitness of agricultural populations vs. non-invasive native populations within an agricultural setting. In this study, I conducted a common garden experiment to test the hypothesis that agricultural weed strains have undergone adaptive evolution for the agricultural environment. My experiment consisted of planting waterhemp sampled from populations across the species range into agricultural field plots in Missouri and Ohio; I replicated the common garden inside (MO) and outside (OH) the area of agriculturally-problematic waterhemp, to test for possible local adaptation to environmental conditions in my source populations. The field plots were planted with soybeans, because waterhemp is the most problematic for this crop in the Midwest. By excluding herbicides, I focused on morphological and life history traits across the growing season to assess the relative fitness of waterhemp from different geographical regions.

I asked two questions: (1) Is "Mississippi Valley" waterhemp from the most heavily invaded states (MO, IL, and IA) more fit in soybean fields than waterhemp from regions with less or no agricultural waterhemp?; and if so, (2) Does agricultural adaptation or local adaptation to Mississippi Valley environments explain the higher fitness? I hypothesized that Mississippi Valley waterhemp plants are better adapted to agricultural environments, and that they are especially adapted to local Mississippi Valley environments. From these hypotheses, I predicted that Mississippi Valley plants would have higher relative fitness than plants from other regions in both MO and OH common gardens.

MATERIALS AND METHODS

For the purposes of the common garden experiments, the geographical range of *Amaranthus tuberculatus* was divided into three regions, which were hypothesized to have varying levels of adaptation to agricultural environments: the Plains region (including KS, NE, and OK populations in the experiment), the Mississippi Valley region (including MO, IA, and IL), and the Northeastern region (including OH, MI, and ON). Agricultural waterhemp is a serious, economically important pest in Missouri, Iowa, and Illinois (and to a lesser extent, in Indiana and western Ohio), whereas it is an opportunistic weed in the Plains states, and not known to occur agriculturally in most of Ohio and farther north. On the basis of these observations, I designed the hypotheses stated in the introduction.

Common Garden Design

In the fall of 2009 and 2010, seeds were collected from populations across the range of *A. tuberculatus* s.l. for use in common garden experiments in two locations: Missouri (summer 2010) and Ohio (summer 2011). The experiment was replicated in Ohio to control for the possibility that the superior performance of Mississippi Valley plants in Missouri common gardens was due to local adaptation, rather than agricultural adaptation. For both common garden experiments, six populations from the three different geographic regions (see above) were selected for the experiment: for consistency between regions, almost all populations were "natural," i.e., not collected in agricultural fields, with the exception of one IA population (Iowa 1), and one IL population (KNK). These two agricultural populations were chosen to maximize geographical spread of sampling, and on the basis of little to no population structure in the species at a state level (KW, unpublished data). Seeds from the two Iowa populations were

obtained from the USDA GRIN database (originally collected in 1989 and 1996 by Donald Pratt), and thus the number of original plants contributing to the seed collection was unknown. Exactly the same populations were used for the Plains region and the Mississippi Valley region between years, and only two substitutions were made for the Northeastern region between years, as two newly-collected Ohio populations were included to sample the "agricultural waterhemp area" of the state, and to correct for possible confounding of latitude of origin with agricultural adaptation. See Table 3.1 for population names and geographical locations.

Seeds from 10 female plants per population were stored at room temperature with silica gel or frozen at -20ºC until three to four months before the common garden experiments. At this point, 16 seeds were randomly selected from each individual (parent) and placed on a damp paper towel inside a ziploc bag, which was labeled with the parent's population and number in that population. These bags were stored at 4ºC for three to four months, and checked every other week to remove decaying seeds and/or change paper towels if mold had started to grow. This stratification procedure mimics the natural winter stratification of shallowly-buried seeds in *A. tuberculatus* habitats. If 16 seeds were not available for a particular parent, as many as were available were used, and supplementary seeds were stratified from another individual in the population. Lack of sufficient seeds/parents was also the rationale for combining several of the Northeastern geographical populations for the common garden in both years (see Table 3.1). Seeds for the 2010 garden were stratified from Feb. 1 to May 18, and seeds from the 2011 garden were stratified from Feb. 14 to June 9. Planting was timed to coincide with soybean planting in both years.

Eight randomly-selected seeds per parent were planted in 98-well flats, with two seeds/labeled well, in the Washington University greenhouse. The newly-planted seeds were

placed on a mist bench for 1-2 days to facilitate germination, and then removed to a warm sunny bench in the greenhouse. The plants were thinned to one per well soon after germination, and poor germination for seeds from a particular parent was compensated with seedlings from another parent from the same population. The position of the seedling flats was randomized on the greenhouse bench every week until transplantation into the common garden. Just prior to transplanting, seedlings were randomly assigned a number from 1-720 (generated in Excel), and the seedlings were arranged in numerical order in sets of 180, for each plot (block) in the common garden. The height of each seedling was also recorded just prior to transplanting, to use as a control for maternal effects.

Between waterhemp stratification and transplanting, the common garden areas were prepared. In 2010, three old field sites at Washington University's Tyson Research Center (Eureka, MO) were chosen on the basis of their similarity and suitability for soybean plots. Plots measured 7 x 10m, and were tilled with a rotary cultivator on May 5. RoundUp Ready soybeans (Asgrow RR3830, Monsanto, St. Louis, Missouri, USA) were planted shallowly in rows by hand between May 19-26, with 19 rows/plot spaced 0.5m apart, and ~150 soybeans/row (4-5 cm apart), according to recommendations found in UM-Extension publications (Helsel and Minor, 1993). Several commercial products were used to deter deer, including Liquid Fence Deer and Rabbit Repellent (The Liquid Fence Co., Blakeslee, Pennsylvania, USA) and Alaska Fish Fertilizer (Alaska Fish Fertilizer Co., Renton, Washington, USA), sprayed directly on the soybean plants, and polypropylene deer fencing was placed around each plot on June 15-16. The waterhemp plants were transplanted into the soybean plots from June 16 to 19, and immediately watered for establishment. The waterhemp rows were 5 meters long (1 m from the fence on either side horizontally and 1.5 m from the fence vertically), and placed between

soybean rows (20 cm from each soybean row and 50 cm from the next waterhemp row), with individuals spaced 40 cm apart in the row to avoid intraspecific competition. Each plot had 13 rows of waterhemp with 13 plants/row, and a 14th row with 11 plants. The plots were hand weeded throughout the growing season to remove all plants other than soybeans and waterhemp.

In 2011, common gardens were located at Miami University's Ecology Research Center (ERC), in Oxford, OH. The 2010 experiment was replicated as nearly as possible in Ohio. The three Oxford plots were randomly placed in a single 27 x 92 m soybean field, for most convenient mechanical soybean planting. Again, plots measured 7 x 10 m, and were tilled and drill-planted with RoundUp Ready soybeans (Genuity Star RR3404, Monsanto, St. Louis, Missouri, USA) on June 8, in rows 20 inches apart. Because of an unusually wet spring in Ohio, soybean planting was delayed compared to the previous year. Deer herbivory on the emerging soybeans was severe, necessitating soybean replanting on June 25. Because of this, waterhemp transplanting was delayed and thus waterhemp seedlings were kept in the Washington University greenhouse longer than in 2010, making it necessary to move the seedlings into a larger pot size (24-well flats) on June 27 to prevent them from becoming pot-bound and stunted. The waterhemp seedlings were transplanted into the common garden plots on July 6-8, and immediately watered for establishment. A four-wire electric fence was put around the entire field on June 27, and polypropylene deer fencing was put up around each plot individually on July 8. The spatial positioning of waterhemp rows and individuals was the same as for the Missouri plots, except that the double soybean planting led to essentially random spacing of soybeans with respect to waterhemp rows. The plots were hand weeded throughout the growing season, and commercial deer repellents including Liquid Fence and DeerOff (Woodstream Co., Lititz, Pennsylvania, USA) were used to deter deer.

Plant Measurements, 2010

As described above, waterhemp seedling height was recorded just before transplanting into the plots, to use as a control for maternal effects. Starting a few days after transplanting and every week thereafter, plant survivorship in the Missouri common garden plots was recorded. Flowering started on June 29, and flowering start date, flowering plant height, and sex of the plant was recorded from June 29 to August 19, every 5-9 days. An open flower on a male or female plant was taken as the start of flowering. Mature plant measurements were taken when the majority of flowers were open (for male plants) or the majority of flowers had set seed (for female plants). Mature height, number of branches off the main stem, and length of longest primary branch were recorded for each waterhemp plant. These measurements were taken between August 13 and October 5, approximately every 2-3 weeks (except Plot 1, for which measurements were recorded on August 13, 19, 29, and October 2).

Immediately after final measurements were taken, the plant's above-ground biomass was removed near the ground and put in a brown paper bag to dry. These bags were stored at Washington University in a dry room containing mothballs to deter insect activity. Because the harvested plants were quite bulky, standard oven or incubator drying for biomass measurements was impossible; therefore, a Conviron plant growth chamber (PGW36 model, Conviron, Winnipeg, Manitoba, Canada) in the Washington University greenhouse was used to dry the plants. The growth chamber was set at 37-39ºC, 20-26% relative humidity, and lamps at 44 watts/m^2. Batches of bags were left in the drier for 9-15 days, at which time each bagged plant was weighed on an electronic scale (preliminary experiments established that dry weight stabilized at 9 days). Dried above-ground biomass measurements were recorded to the nearest

0.01 gram. Five empty bags of each size were also dried for at least 9 days, and the average weight of each size bag was subtracted from the biomass of plants in that size bag. The number or weight of seeds per plant was not quantified, because waterhemp plants in crop fields can produce up to a million seeds each (Steckel, 2007), and seed production is not confined to discrete seed heads but spread over the entire plant. Thus, seed production was judged impractical to measure, and dry above-ground biomass was taken to be the most feasible measure of fitness.

Plant Measurements, 2011

The same procedures were followed for the Ohio common garden, except for modifications described below. Because the region around Oxford, OH does not yet have a problem with agricultural waterhemp, procedures were implemented to contain gene flow from the experimental waterhemp into surrounding agricultural fields and/or nearby riverbank populations. Therefore, waterhemp in the Ohio common garden was monitored much more frequently than in the Missouri common garden: survival, flowering start date, and flowering plant height were recorded every 2-3 days, from July 11-August 19. Furthermore, to prevent the pollen from being dispersed by wind, male plants were measured for mature data and harvested as soon as their first flower opened (therefore, male plant flowering height and mature height were the same). As males grow very little after flowering begins, this difference is unlikely to have influenced inferences (see Results). Female plants were measured for final data and harvested at approximately the same point as in the Missouri common garden, before many seeds/fruits could drop from the plant. Male plants were harvested every two-three days from June 15 to September 2. Female plants were harvested every 2-3 weeks from September 2 to

October 12. The bagged plants were stored in a trailer at the ERC until they could be brought back to Washington University and stored in the same room as the Missouri bagged plants. Dried biomass for these plants was measured in exactly the same way as for the Missouri plots.

Data Analysis

All plant measurement data were analyzed using PASW Statistics 18.0.0 for Windows (SPSS Inc., Hong Kong, China). First, all continuous data were tested for normality using the Shapiro-Wilk test. If the data were not normal, they were either $\log_{10}$ transformed or square-root transformed. If data were still not normal after transformation (e.g., flowering plant height in the 2010 plots), nonparametric tests were used for these data, as well as for ordinal data (days to flowering).

A univariate general linear model (GLM) or the equivalent nonparametric test was used to analyze most data, including height at transplantation, flowering height, mature height, branch number, length of longest branch, flowering start date, and dry above-ground biomass. Height at transplantation was subtracted from subsequent height measurements to control for maternal effects. A repeated-measures general linear model was also used to analyze height over time, and a multivariate GLM was used to analyze mature plant height, branch number, and longest branch length together, because of the non-independence of these measurements. The fixed factor in each GLM was geographical region of origin (Plains, Mississippi Valley, Northeast), with block (plot) and population nested within region as random factors. For significant results, post-hoc Tukey HSD tests were used to determine whether means were significantly different between each pair of regions.

For the Missouri plots, data were analyzed with and without the inclusion of plants that started to show inflorescence development before transplantation, or that were dead when mature measurements were taken (due to the spacing between mature data collection points). For the Ohio plots, data were analyzed with and without plants infested with ash-gray leaf bugs (*Piesma cinerea* Say) (which appeared to have stunted plant growth). For both years, days to flowering, flowering plant height, mature plant data, and dry above-ground biomass were also analyzed by plant sex and by genetic subpopulation (described below). Finally, to rule out any confounding factors introduced by harvesting the Ohio male plants earlier than the Missouri males, only female data were analyzed for both plots and compared to the full data set.

RESULTS

Mortality after establishment and before maturity was very low in both years. In total, 629 of 720 plants survived the transplantation period in 2010, and 14 of these established plants died during the growing season. Mortality in 2010 stemmed almost entirely from a "damping off" fungal infection that killed the plants within 10 days of transplantation, without regard for geographic region of origin. In 2011, 699 plants out of 720 survived transplantation, and only one of these survivors died during the growing season. In both years, analyses with and without early-blooming/early-dying/damaged plants (see Methods) had generally consistent results, with lower significance for the datasets with these plants removed (probably because of lower sample sizes); to be conservative, results for the latter datasets are reported below.

Plant Height

There was a significant effect of region on height at transplantation in 2011 ($F_{2,15.01}$=3.538, P=0.041) and a marginally significant effect in 2010 ($F_{2,15.01}$=3.991, P=0.055;

Table 3.2). Posthoc Tukey HSD tests showed that plants from the Northeastern region were on average shorter than plants from the other two regions ($P_{2,3}$ and $P_{1,3}<0.001$ for 2010 and 2011). The magnitude of this regional height difference increased as the plants grew, with Northeastern plants' average flowering height being significantly shorter in 2010 (Kruskal-Wallis test, df=2, $\chi^2=23.480$, $P<0.001$) and 2011 ($F_{2,15.07}=6.365$, $P=0.01$; Posthoc Tukey HSD tests, $P_{2,3}$ and $P_{1,3}$ <0.001 for both years). Mature Northeastern plants were also significantly shorter in both years (2010, $F_{2,15.81}=12.565$, $P=0.001$; 2011, $F_{2,15.11}=6.073$, $P=0.012$). Mississippi Valley plants were the tallest at maturity in 2010 (Posthoc Tukey HSD tests, $P_{1,2}=0.001$, $P_{2,3}$ and $P_{1,3}<0.001$), but there was no difference between mature heights of Mississippi Valley and Plains plants in 2011 ($P_{1,2}= 0.800$, $P_{2,3}$ and $P_{1,3}$ <0.001). Repeated measures analyses of longitudinal height data (transplant, flowering, and mature) showed that region of origin was highly significant over time, and that Northeastern plants were always shorter on average (2010: $F_{4,518}=20.624$, $P<0.001$; 2011: $F_{4,762}=31.552$, $P<0.001$; Posthoc Tukey HSD tests, $P_{2,3}$ and $P_{1,3}<0.001$ both years; Figure 3.2a, 3.2b). Block and population (nested within region) were also significant for all of these analyses (Figure 3.3a, 3.3b). Plant height over time provides an approximate measure of growth rate, and these data suggest that Northeastern plants grow more slowly (and mature at smaller stature) than plants from the other regions.

Days to Flowering

Kruskal-Wallis tests of the ordinal data "days to flower," the number of days from planting to flowering, showed that Northeastern plants flowered significantly earlier in both years (2010, df=2, $\chi^2=12.237$, $P=0.002$, Northeastern plants flowering an average of 6.56 days earlier than the other two regions averaged together; 2011, df=2, $\chi^2=11.542$, $P=0.003$,

Northeastern plants flowering 3.78 days earlier on average) (Table 3.2; Figure 3.4a, 3.4b). This effect was influenced strongly by latitude of origin in 2010, with four out of six Northeastern populations coming from relatively high latitude sites. Nonetheless, when two of these high-latitude populations were replaced with lower-latitude Northeastern populations in 2011, flowering date was still significantly earlier for this region. This suggests that Northeastern plants respond differently to photoperiod cues than do Plains or Mississippi Valley plants and that this difference is not solely attributable to differences in latitude of origin.

Mature Plant Data

To estimate of the size of mature plants, three measurements were taken just before plants were harvested: height, branch number, and length of the longest branch. In a multivariate analysis, region of origin made a significant difference in 2010 mature height ($F_{2,257}$=39.844, P<0.001) and branch number ($F_{2,257}$=15.565, P<0.001), but not length of the longest branch ($F_{2,257}$=0.992, P=0.372; Table 3.2). Posthoc Tukey HSD tests showed that Mississippi Valley plants were the tallest ($P_{1,2}$=0.001, $P_{2,3}$ and $P_{1,3}$<0.001) and had the most branches ($P_{1,2}$=0.002, $P_{2,3}$<0.001; Figure 3.5a). In the 2011 multivariate analysis of mature data, region of origin has a significant effect on mature height ($F_{2,379}$=52.104, P<0.001), branch number ($F_{2,379}$=6.442, P=0.002), and length of longest branch ($F_{2,379}$=21.535, P<0.001). Mississippi Valley and Plains plants were on average taller and had longer longest branches than Northeastern plants ($P_{2,3}$ and $P_{1,3}$<0.001 for both); but Plains plants had the fewest average branches ($P_{1,2}$=0.002; Figure 3.5b). Taken together, these results suggest that Northeastern plants are always smaller overall at maturity, but that the relative average fitness of Mississippi Valley and Plains individuals depends on common garden location.

Dry Above-Ground Biomass

The impact of region of origin on dry biomass was significant in 2010 ($F_{2,15.84}$=5.809, P=0.013) and in 2011 (Kruskal-Wallis test, df=2, χ^2=17.516, P<0.001; Table 3.2; Figure 3.6a, 3.6b). In 2010, Mississippi Valley plants were heavier on average than plants from the other two regions (Posthoc Tukey HSD test, $P_{1,2}$=0.001, $P_{2,3}$<0.001); in 2011, Mississippi Valley and Plains plants were heavier on average than Northeastern plants ($P_{2,3}$ and $P_{1,3}$<0.001). Again, the relative fitness of plants from the Mississippi Valley and Plains regions appears to depend on common garden location, whereas the Northeastern plants appear to have the lowest average fitness regardless.

Analyses by Plant Sex

When data were analyzed with sex included in the GLM with region, female plants were consistently taller, heavier, and later flowering, and they had more branches (and longer longest branches in 2011) than male plants did, regardless of region or population (Table 3.3; Figures 3.7-3.10). Because male plants were measured for mature data and harvested earlier in 2011 than in 2010, the female plant data from both years were also analyzed separately. With female data alone, 2010 and 2011 analyses showed the same patterns as with data from all plants, but the relationships were not as highly significant (Table 3.4). Days to flowering was still significantly different between regions in the 2010 experimental plots (df=2, χ^2=6.004, P=0.050, with Northeastern plants flowering earlier). Flowering height ($F_{2,15.38}$=4.450, P=0.030), mature height ($F_{2,128}$=27.373, P<0.001), branch number ($F_{2,128}$=11.790, P<0.001), and dry biomass ($F_{2,17.20}$=5.000, P=0.019) were all significantly different between regions, with Mississippi

Valley plants heavier and with more branches, and Northeastern plants shorter on average. Length of the longest branch was again not significantly affected by region in 2010 ($F_{2,128}$=0.529, P=0.590). In 2011, days to flowering (df=2, χ^2=9.295, P=0.010), flowering height ($F_{2,17.68}$=3.939, P=0.038), mature height ($F_{2,138}$=19.330, P<0.001), branch number ($F_{2,138}$=12.204, P<0.001), length of longest branch ($F_{2,138}$=3.494, P=0.033), and dry biomass ($F_{2,19.63}$=4.035, P=0.034) were all significantly different between regions, although Plains female plants flowered significantly earlier than the other two regions' females. Together, these analyses indicate that omitting male plant data yields generally the same results for both years; thus, harvesting males earlier in 2011 did not have a significant effect on the overall results.

DISCUSSION

I conducted a common garden experiment in two different geographical locations, one inside and one outside the range of agriculturally-invasive waterhemp, to determine whether waterhemp plants from the Mississippi Valley region have higher fitness in agricultural environments than do plants from regions without problematic agricultural waterhemp. I also tested whether these fitness relationships hold regardless of geographical location. I found that regardless of location and year, seeds derived from Northeastern plants (from OH, MI, and ON) are less fit in soybean plots than are seeds from Plains and Mississippi Valley source populations. On average, Northeastern plants grow more slowly and reach a smaller maximum size and weight than do plants from the other two regions. Interestingly, Northeastern plants were significantly shorter even before transplantation into soybean plots. This phenotypic difference is likely to play a key role in Northeastern plants' lack of competitiveness in the agricultural environment. These results allow me to conclude that there are adaptive differences

between populations from different geographical regions, and that these differences contribute to the differential fitness of Mississippi Valley plants as agricultural invasives.

Northeastern plants also flowered between 3 to 7 days earlier on average than did plants from the other two regions. *Amaranthus tuberculatus* s.l. is a short-day plant (Costea et al., 2005), although photoperiod is only one of the factors controlling flowering time (e.g., plants will flower when very small if pot-bound [K. Waselkov, pers. obs.]). My field observations suggest that crop field waterhemp populations typically flower earlier than do nearby riverbank populations in the agricultural waterhemp regions, despite the near-certainty of high gene flow between these populations; this suggests that waterhemp flowering phenology responds plastically to agricultural practices. Flowering is the beginning of senescence for waterhemp individuals, particularly for males: thus, earlier flowering limits the size that Northeastern waterhemp can attain during the growing season. Life-history events such as flowering are phenotypically plastic traits under strong selection in crop fields, as a weed's growing season is entirely bounded by crop planting and harvest (Ghersa and Holt, 1995; Neve et al., 2009). One of the few authors to address the importance of phenology to agricultural weeds was Barrett (1983), who observed that *Echinochloa crus-gallii* crop mimics had evolved to match the phenology of the crop. More broadly, examination of life-history traits such as flowering time, fecundity, and dormancy suggests that variation in agricultural practices can select for different life-history strategies in a single species, as observed in *Capsella bursa-pastoris* in the UK (Begg et al., 2012).

Contrary to the consistent results between years for Northeastern plants, plants from the Mississippi Valley and Plains regions showed different patterns depending on the year/location of the common garden. In the 2010 experiment conducted in Missouri, Mississippi Valley plants

outperformed plants from the other regions in mature height, number of branches, and dry biomass measurements. In the 2011 Ohio experiment, Mississippi Valley and Plains plants were not significantly different for mature height, length of longest branch, or dry biomass measurements, although Plains plants had fewer branches on average. These patterns provide evidence for local adaptation of Mississippi Valley plants, with source populations from MO, IA, and IL. These plants may have been better suited to the soil and climatic conditions of the Eureka, MO common garden than those of the Oxford, OH common garden, which is located slightly outside the zone of agricultural waterhemp infestation. I cannot rule out the possibility that the later planting date and randomly-placed soybeans in the Ohio plots might have contributed to the difference in results between years: the only way to disentangle these factors would be to perform more years of common garden studies to control for inevitable climate and pest variation between any two years of outdoor research. However, from the generally similar results derived from the analysis of females alone versus both sexes for each year, I can conclude that harvesting male plants earlier in 2011 than in 2010 did not significantly change the trends in the data.

In 2010, a simultaneous population genetics study with 10 microsatellite markers was conducted using populations from across the species' range. This study revealed two genetic subpopulations within waterhemp, broadly divided by the Mississippi River, but with some populations east of the Mississippi (largely confined to the "agricultural waterhemp" regions of Illinois, Indiana, and Ohio) showing genetic affinity to the western group (K. Waselkov, in prep.). When the populations used in the common garden experiments are considered based on their genetic subpopulation, the Plains and Mississippi Valley populations almost all fall within

the "western" genetic group, with the exception of KNK and KEY populations from IL, which fall into the "eastern" genetic group with all of the Northeastern populations (Table 3.1).

When the common garden data from 2010 and 2011 are analyzed by genetic subpopulation, rather than geographical region, patterns of lower "eastern" fitness emerge: days to flowering, flowering height, mature height, and longitudinal height analyses all show significant effects of genetic region in both years. Plants from populations in the "eastern" genetic subpopulation were shorter and flowered earlier on average than did plants from the "western" subpopulation (Table 3.5). Branch number is significantly affected by genetic region in 2010 (GLM: $F_{1,257}$=6.808, P=0.010), but not in 2011 ($F_{1,379}$=0.606, P=0.437), and length of longest branch is significantly affected by genetic region in 2011 ($F_{1,379}$=34.819, P<0.001), but not in 2010 ($F_{1,257}$=1,187, P=0.277). Dry biomass was significantly less for genetically "eastern" plants in 2011 (Mann-Whitney U test, df=1, Z=-3.599, P<0.001), but not in 2010 ($F_{1,17.76}$=1.650, P=0.215) (Figure 3.11). These results demonstrate that the "genetic subpopulation" division of source populations gives results congruent with division by areas of agricultural infestation.

Although not conclusive, the combination of my common garden fitness data and genetic results strongly suggests that Mississippi Valley and/or Plains populations were "preadapted" to invade Mississippi Valley agricultural environments when the opportunity presented itself in the 20th century, rather than requiring genetic changes to become successful in these new habitats. The genetic similarity between Mississippi Valley and Plains source populations (despite the different levels of agricultural infestation in these regions) and dissimilarity from the Northeastern source populations indicate that the "western" genetic variety may have already possessed the qualities necessary to compete with crops (K. Waselkov, in prep). However, from my experiments, I cannot pinpoint which morphological or life-history traits or environmental

variables in particular lead to higher fitness in the Mississippi Valley/Plains plants. This would require multifactorial common gardens or controlled greenhouse experiments.

The question of agricultural "preadaptation" has seldom been addressed, because few weeds have invaded agricultural environments recently enough to permit examination of "before and after" populations. Waterhemp is unusual in that the approximate time and location of its agricultural invasion are known. In invasion biology, there is much interest in predicting invasiveness based on particular morphological or life-history traits (Kolar and Lodge, 2001), and several researchers have taken advantage of knowing the details of recent invasions to compare these traits in conspecific invasive and native populations (e.g., Leger and Rice, 2003; Erfmeier and Bruelheide, 2005; Caño et al., 2008). In general, these studies have shown greater fitness of the invasive populations, suggesting genetic adaptation rather than preadaptation. Other researchers have addressed preadaptation in a different way, by comparing the growth of species that have and have not invaded other continents in common gardens in their native range, with seeds from native populations (Schlaepfer et al., 2010; Van Kleunen et al., 2011). Contrary to the conspecific experiments, these interspecific studies provide evidence for pre-adaptation of invasive species through species traits that confer higher fitness in their native range, such as high biomass production and fast growth rate.

In contrast to Baker's 1974 "general-purpose genotype" hypothesis, which proposed that phenotypic plasticity alone can explain most weed adaptiveness, a consensus is developing among weed scientists that evolution should be taken into account when developing integrated pest-management strategies (Clements et al., 2004; Neve et al., 2009). Evolutionary biologists are also starting to take more interest in agricultural weeds, as they often exhibit microevolution on ecological time scales, driven by strong (albeit unintentional) anthropogenic selection

(Vigueira et al., 2013). Agricultural environments have never been static, and changing cultivation practices in the late 20th century have provided an opportunity to observe rapid evolutionary change in weeds. With the intensification of farming since the 1940s, herbicide resistance has evolved in 216 species worldwide since 1978, with an average of nine new cases of resistance emerging per year (Heap, 2013). Conservation tillage has also changed the species that are most problematic agriculturally (Swanton et al., 1993; Buhler, 1995). Furthermore, the introduction of glyphosate-resistant crops in the 1990s, which by 2006 made up 89% of soybean and 39% of corn production in the U.S., led to both increased reliance on herbicide and reduced tilling, and subsequently to shifts in the weed community in these fields (Hawes et al., 2003; Hilgenfeld et al., 2004; Owen, 2008).

Thus far, explicitly evolutionary studies of agricultural weeds have most often focused on discovering their origin (e.g., Burger et al., 2007; Muller et al., 2010; Reagon et al., 2010), and weeds related to crops have taken precedence due to the possibilities for hybridization and introgression between crops and their weedy relatives (e.g. Bartsch et al., 2003; Warwick et al., 2003; Snow et al., 2010; reviewed in Ellstrand et al., 1999). Evolution of herbicide resistance is an exception to these research patterns, being of great interest to both weed scientists and evolutionary biologists (reviewed in Jasieniuk et al., 1996). Other than herbicide resistance (often a single-gene or even single-nucleotide trait), intraspecific variation in agricultural adaptation of weedy plants has been understudied (but see Mercer et al., 2002; Lai et al., 2008; Begg et al., 2012).

Most notable common garden studies of adaptive evolution in agricultural weeds are European: Keller et al. (2000) found evidence for local adaptation in Switzerland of three arable weed species by comparing the fitness of parents and outcrossed progeny in a common garden.

In Sweden, Bommarco et al. (2010) detected common garden differences in competitive ability between thistles (*Cirsium arvense*) from agricultural, ruderal, and natural habitats. Neither of these studies grew their focal weeds in an agricultural setting for the common garden experiment. In contrast, Leiss and Müller-Scharer (2001) performed reciprocal transplant experiments between ruderal and agricultural habitats for *Senecio vulgaris* in Switzerland, and found no evidence for local adaptation. To the best of my knowledge, only one experiment of this type has previously been conducted in the U.S. with a native weed species: Hartnett et al. (1987) reciprocally transplanted *Ambrosia trifida* between agricultural habitats of different successional stages in Illinois, finding adaptive differences between these two populations but no evidence for local adaptation.

My study differs from previous experiments with waterhemp in agricultural plots in that all of these studies focused on the control of waterhemp for crop production. Several previous experiments testing waterhemp fitness in agricultural environments focused on the impact of waterhemp on soybean yield (Hager et al., 2002b; Steckel and Sprague, 2004). Other waterhemp common garden experiments were explicitly designed to test the best methods for controlling waterhemp with herbicides (Steckel et al., 2002; Legleiter et al., 2009). Still others specifically measured waterhemp fitness as a function of emergence date in agricultural plots, with the goal of determining when best to control waterhemp (Hartzler et al., 2004; Nordby and Hartzler, 2004). No previous study has compared seeds from natural populations of waterhemp that were hypothesized to have different levels of agricultural adaptation.

My experiments were not designed to measure seed dormancy and germination life-history traits, which several previous publications have shown to vary among waterhemp populations and among tillage systems (Leon and Owen, 2006; Leon et al., 2007; Refsell and

Hartzler, 2009). A particularly interesting study showed, albeit with single individuals, that an Ohio riverbank plant had much lower seed dormancy than did two Iowa agricultural plants (Leon et al., 2006). The present study minimized the impact of seed dormancy differences on fitness by stratifying all the seeds, which made germination more even between Ohio and Iowa plants in Leon et al.'s 2006 study. Future common garden experiments with natural waterhemp populations should aim to incorporate seed dormancy characteristics, as these traits have a large impact on fitness in other agricultural weeds (Benech-Arnold et al., 2000; Shivrain et al., 2009; Norsworthy et al., 2010).

Finally, common gardens are the most basic type of experiment for studies of local adaptation and intraspecific trait variation. For invasive species, researchers should ideally have common gardens in both the native and introduced range, as I do here, to control for any interactions between the garden location and the genetic provenances of the plants (Hierro et al., 2005; Moloney et al., 2009). Reciprocal transplant experiments are even more sophisticated, as they measure the performance of plants in each other's native environments (Kawecki and Ebert, 2004). Therefore, the ideal waterhemp garden experiment would be reciprocal transplants of waterhemp from the agriculturally invaded and uninvaded ranges into both soybean plots and riverbank plots. Unfortunately, problems with extensive riverbank flooding in 2010 prohibited transplantation of waterhemp into riverbank plots in Missouri (as originally planned). However, for future studies, paired, replicated riverbank and crop field plots, in several sites inside and outside the range of agricultural waterhemp, would be the most comprehensive way to study fitness and local adaptation in this system. These experiments would shed further light on whether the small size of Northeastern waterhemp is adaptive in the environments where it naturally occurs. They could also be designed to test for fitness tradeoffs resulting from

herbicide resistance, which have seldom been documented in *Amaranthus* species for resistance to herbicides other than atrazine (Sibony and Rubin, 2002; Gassmann, 2005; Duff et al., 2009; but see Tardif et al., 2005).

I found evidence that *Amaranthus tuberculatus* s.l. from the geographical region where the species is agriculturally invasive is better adapted to crop field environments than are plants from populations outside the agriculturally invaded region, and that waterhemp from the most heavily invaded region is also locally adapted to Mississippi Valley environments. These results have implications for the evolution of new native agricultural weeds, particularly the evidence for "preadaptation" of a subset of *A. tuberculatus* s.l. to crop fields: many species in naturally disturbed environments like riverbanks may already have traits that would confer high fitness in agricultural environments, and their invasion could be precipitated by changes in management practices (such as conservation tillage and reliance on herbicide, in the case of waterhemp). This study also has implications for future studies of rapid evolution in plants. My results are the latest in a growing body of evidence that evolutionary factors, such as population structure, adaptive genetic variation, and response to selection, are important in shaping invasiveness in agricultural weeds, as well as invaders of more natural ecosystems.

LITERATURE CITED—

AYRES, D. R., D. L. SMITH, K. ZAREMBA, S. KLOHR, AND D. R. STRONG. 2004. Spread of exotic cordgrasses and hybrids (*Spartina* sp.) in the tidal marshes of San Francisco Bay, California, USA. *Biological Invasions* 6: 221-231.

BAKER, H. G. 1974. The evolution of weeds. *Annual Review of Ecology and Systematics* 5: 1-24.

BARRETT, S. C. H. 1983. Crop mimicry in weeds. *Economic Botany* 37: 255-282.

BARTSCH, D., J. CUGUEN, E. BIANCARDI, AND J. SWEET. 2003. Environmental implications of gene flow from sugar beet to wild beet – current status and future research needs. *Environmental Biosafety Research* 2: 1-11.

BEGG, G. S., J. WISHART, M. W. YOUNG, G. R. SQUIRE, AND P. P. M. IANNETTA. 2012. Genetic structure among arable populations of *Capsella bursa-pastoris* is linked to functional traits and in-field conditions. *Ecography* 35: 446-457.

BENECH-ARNOLD, R. L., R. A. SÁNCHEZ, F. FORCELLA, B. C. KRUK, AND C. M. GHERSA. 2000. Environmental control of dormancy in weed seed banks in soil. *Field Crops Research* 67: 105-122.

BOMMARCO, R., M. LÖNN, U. DANZER, K-J. PÅLSSON, AND P. TORSTENSSON. 2010. Genetic and phenotypic differences between thistle populations in response to habitat and weed management practices. *Biological Journal of the Linnean Society* 99: 797-807.

BUHLER, D. 1995. Influence of tillage systems on weed population dynamics and management in corn and soybean in the Central USA. *Crop Science* 35: 1247-1258.

BURGER, J. C., J. M. HOLT, AND N. C. ELLSTRAND. 2007. Rapid phenotypic divergence of feral rye from domesticated cereal rye. *Weed Science* 55: 204-211.

CAÑO, L., J. ESCARRÉ, I. FLECK, J. M. BLANCO-MORENO, AND F. X. SANS. 2008. Increased fitness and plasticity of an invasive species in its introduced range: a study using *Senecio pterophorus*. *Journal of Ecology* 96: 468-476.

CLEMENTS, D. R., A. DITOMMASO, N. JORDAN, B. D. BOOTH, J. CARDINA, D. DOOHAN, C. L. MOHLER, ET AL. 2004. Adaptability of plants invading North American cropland. *Agriculture, Ecosystems and Environment* 104: 379-398.

COSTEA, M., AND F. J. TARDIF. 2003. Conspectus and notes on the genus *Amaranthus* in Canada. *Rhodora* 105: 260-281.

COSTEA, M., S. E. WEAVER, AND F. TARDIF. 2005. The biology of invasive alien plants in Canada. 3. *Amaranthus tuberculatus* (Moq.) Sauer var. *rudis* (Sauer) Costea & Tardif. *Canadian Journal of Plant Science* 85: 507-522.

DAVIS, V. M., G. R. KRUGER, J. M. STACHLER, M. M. LOUX, AND W. G. JOHNSON. 2009. Growth and seed production of horseweed (*Conyza canadensis*) populations resistant to

glyphosate, ALS-inhibiting, and multiple (glyphosate + ALS-inhibiting) herbicides. *Weed Science* 57: 494-504.

DE WET, J. M. J., AND J. R. HARLAN. 1975. Weeds and domesticates: evolution in the man-made habitat. *Economic Botany* 29: 99-107.

DÉLYE, C., S. MICHEL, A. BÉRARD, B. CHAUVEL, D. BRUNEL, J-P. GUILLEMIN, F. DESSAINT, AND V. LE CORRE. 2010. Geographical variation in resistance to acetyl-coenzyme A carboxylase-inhibiting herbicides across the range of the arable weed *Alopecurus myosuroides* (black-grass). *New Phytologist* 186: 1005-1017.

DLUGOSCH, K. M., AND I. M. PARKER. 2008. Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology* 17: 431-449.

DUFF, M. G., K. AL-KHATIB, AND D. E. PETERSON. 2009. Relative competitiveness of protoporphyrinogen oxidase-resistant common waterhemp (*Amaranthus rudis*). *Weed Science* 57: 169-174.

EBELING, S. K., J. STÖCKLIN, I. HENSEN, AND H. AUGE. 2011. Multiple common garden experiments suggest lack of local adaptation in an invasive ornamental plant. *Journal of Plant Ecology* 4: 209-220.

ELLSTRAND, N. C., H. C. PRENTICE, AND J. F. HANCOCK. 1999. Gene flow and introgression from domesticated plants into their wild relatives. *Annual Review of Ecology and Systematics* 30: 539-563.

ERIKSEN, R. L., T. DESRONVIL, J. L. HIERRO, AND R. KESSELI. 2012. Morphological differentiation in a common garden experiment among native and non-native specimens of the invasive weed yellow starthistle (*Centaurea solstitialis*). *Biological Invasions* 14: 1459-1467.

ERFMEIER, A., AND H. BRUELHEIDE. 2005. Invasive and native *Rhododendrom ponticum* populations: is there evidence for genotypic differences in germination and growth? *Ecography* 28: 417-428.

GARCÍA-RAMOS, G., AND D. RODRÍGUEZ. 2002. Evolutionary speed of species invasions. *Evolution* 56: 661-668.

GASKIN, J. F., AND B. A. SCHAAL. 2002. Hybrid *Tamarix* widespread in U.S. invasion and undetected in native Asian range. *Proceedings of the National Academy of Sciences* 99: 11256-11259.

GASSMANN, A. J. 2005. Resistance to herbicide and susceptibility to herbivores: environmental variation in the magnitude of an ecological trade-off. *Oecologia* 145: 575-585.

GHERSA, C. M., AND J. S. HOLT. 1995. Using phenology prediction in weed management: a review. *Weed Research* 35: 461-470.

HAGER, A. G., L. M. WAX, G. A. BOLLERO, AND F. W. SIMMONS. 2002a. Common waterhemp (*Amaranthus rudis* Sauer) management with soil-applied herbicides in soybean (*Glycine max* (L.) Merr.). *Crop Protection* 21: 277-283.

HAGER, A. G., L. M. WAX, E. W. STOLLER, AND G. A. BOLLERO. 2002b. Common waterhemp (*Amaranthus rudis*) interference in soybean. *Weed Science* 50: 607-610.

HARLAN, J. R., AND J. M. J. DE WET. 1965. Some thoughts about weeds. *Economic Botany* 19: 16-24.

HARTNETT, D. C., B. B. HARTNETT, AND F. A. BAZZAZ. 1987. Persistence of *Ambrosia trifida* populations in old fields and responses to successional changes. *American Journal of Botany* 74: 1239-1248.

HARTZLER, R. G., B. A. BATTLES, AND D. NORDBY. 2004. Effect of common waterhemp (*Amaranthus rudis*) emergence date on growth and fecundity in soybean. *Weed Science* 52: 242-245.

HAWES, C., A. J. HAUGHTON, J. L. OSBORNE, D. B. ROY, S. J. CLARK, J. N. PERRY, P. ROTHERY, ET AL. 2003. Responses of plants and invertebrate trophic groups to contrasting herbicide regimes in the Farm Scale Evaluations of genetically modified herbicide-tolerant crops. *Philosophical Transactions of the Royal Society of London, B: Biological Sciences* 358: 1899-1913.

HEAP, I. 2013. International survey of herbicide resistant weeds. Website accessed April 10th, 2013. http://www.weedscience.org/summary/home.aspx

HELSEL, Z. R., AND H. C. MINOR. 1993. Soybean production in Missouri. University of Missouri Extension, Department of Agronomy, Publication G4410. Columbia, Missouri, USA.

HIERRO, J. L., J. L. MARON, AND R. M. CALLAWAY. 2005. A biogeographical approach to plant invasions: the importance of studying exotics in their introduced *and* native range. *Journal of Ecology* 93: 5-15.

HILGENFELD, K. L., A. R. MARTIN, D. A. MORTENSEN, AND S. C. MASON. 2004. Weed management in a glyphosate resistant soybean system: weed species shifts. *Weed Technology* 18: 284-291.

HODGINS, K. A., AND L. RIESEBERG. 2011. Genetic differentiation in life-history traits of introduced and native common ragweed (*Ambrosia artemisiifolia*) populations. *Journal of Evolutionary Biology* 24: 2731-2749.

HODGINS, K. A., Z. LAI, K. NURKOWSKI, J. HUANG, AND L. H. RIESEBERG. 2013. The molecular basis of invasiveness: differences in gene expression of native and introduced common ragweed (*Ambrosia artemisiifolia*) in stressful and benign environments. *Molecular Ecology.* doi: 10.1111/mec.12179

HUEY, R. B., G. W. GILCHRIST, M. L. CARLSON, D. BERRIGAN, AND L. SERRA. 2000. Rapid evolution of a geographic cline in size in an introduced fly. *Science* 287: 308-309.

JASIENIUK, M., A. L. BRÛLÉ-BABEL, AND I. N. MORRISON. 1996. The evolution and genetics of herbicide resistance in weeds. *Weed Science* 44: 176-193.

KAWECKI, T. J., AND D. EBERT. 2004. Conceptual issues in local adaptation. *Ecology Letters* 7: 1225-1241.

KELLER, M., J. KOLLMANN, AND P. J. EDWARDS. 2000. Genetic introgression from distant provenances reduces fitness in local weed populations. *Journal of Applied Ecology* 37: 647-659.

KOLAR, C. S., AND D. M. LODGE. 2001. Progress in invasion biology: predicting invaders. *Trends in Ecology and Evolution* 16: 199-204.

KOLBE, J. J., R. E. GLOR, L. RODRÍGUEZ SCHETTINO, A. R. CHEMIZO LARA, A. LARSON, AND J. B. LOSOS. 2004. Genetic variation increases during biological invasion by a Cuban lizard. *Nature* 431: 177-181.

LAI, Z., N. C. KANE, Y. ZOU, AND L. H. RIESEBERG. 2008. Natural variation in gene expression between wild and weedy populations of *Helianthus annuus*. *Genetics* 179: 1881-1890.

LAMBRINOS, J. G. 2004. How interactions between ecology and evolution influence contemporary invasion dynamics. *Ecology* 85: 2061-2070.

LAVERGNE, S., AND J. MOLOFSKY. 2007. Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences* 104: 3883-3888.

LEE, C. E. 2002. Evolutionary genetics of invasive species. *Trends in Ecology and Evolution* 17: 386-391.

LEGER, E. A., AND K. J. RICE. 2003. Invasive California poppies (*Eschscholzia californica* Cham.) grow larger than native individuals under reduced competition. *Ecology Letters* 6: 257-264.

LEGER, E. A., AND K. J. RICE. 2007. Assessing the speed and predictability of local adaptation in invasive California poppies (*Eschscholzia californica*). *Journal of Evolutionary Biology* 20: 1090-1103.

LEGLEITER, T. R., K. W. BRADLEY, AND R. E. MASSEY. 2009. Glyphosate-resistant waterhemp (*Amaranthus rudis*) control and economic returns with herbicide programs in soybean. *Weed Technology* 23: 54-61.

LEISS, K. A., AND H. MÜLLER-SCHARER. 2001. Performance of reciprocally sown populations of *Senecio vulgaris* from ruderal and agricultural habitats. *Oecologia* 128: 210-216.

LEON, R. G., AND M. D. K. OWEN. 2006. Tillage systems and seed dormancy effects on common waterhemp (*Amaranthus tuberculatus*) seedling emergence. *Weed Science* 54: 1037-1044.

LEON, R. G., D. C. BASSHAM, AND M. D. K. OWEN. 2006. Germination and proteome analyses reveal intraspecific variation in seed dormancy regulation in common waterhemp (*Amaranthus tuberculatus*). *Weed Science* 54: 305-315.

LEON, R. G., D. C. BASSHAM, AND M. D. K. OWEN. 2007. Thermal and hormonal regulation of the dormancy-germination transition in *Amaranthus tuberculatus* seeds. *Weed Research* 47: 335-344.

LOOMIS, E. S., AND L. FISHMAN. 2009. A continent-wide clone: population genetic variation of the invasive plant *Hieracium aurantiacum* (orange hawkweed; Asteraceae) in North America. *International Journal of Plant Sciences* 170: 759-765.

MARON, J. L., M. VILÀ, R. BOMMARCO, S. ELMENDORF, AND P. BEARDSLEY. 2004. Rapid evolution of an invasive plant. *Ecological Monographs* 74: 261-280.

MERCER, K. L., N. R. JORDAN, D. L. WYSE, AND R. G. SHAW. 2002. Multivariate differentiation of quackgrass (*Elytrigia repens*) from three farming systems. *Weed Science* 50: 677-685.

MOLONEY, K. A., C. HOLZAPFEL, K. TIELBÖRGER, F. JELTSCH, AND F. M. SCHURR. 2009. Rethinking the common garden in invasion research. *Perspectives in Plant Ecology, Evolution and Systematics* 11: 311-320.

MULLER, M-H., F. DÉLIEUX, J. M. FERNÁNDEZ-MARTÍNEZ, B. GARRIC, V. LECOMTE, G. ANGLADE, M. LEFLON, ET AL. 2009. Occurrence, distribution and distinctive morphological traits of weedy *Helianthus annuus* L. populations in Spain and France. *Genetic Resources and Crop Evolution* 56: 869-877.

MULLER, M.-H., M. LATREILLE, AND C. TOLLON. 2010. The origin and evolution of a recent agricultural weed: population genetic diversity of weedy populations of sunflower (*Helianthus annuus* L.) in Spain and France. *Evolutionary Applications* 4: 499-514.

NEVE, P., M. VILA-AIUB, AND F. ROUX. 2009. Evolutionary-thinking in agricultural weed management. *New Phytologist* 184: 783-793.

NORDBY, D. E., AND R. G. HARTZLER. 2004. Influence of corn on common waterhemp (*Amaranthus rudis*) growth and fecundity. *Weed Science* 52: 255-259.

NORSWORTHY, J. K., M. S. MALIK, M. B. RILEY, AND W. BRIDGES JR. 2010. Time of emergence affects survival and development of wild radish (*Raphanus raphanistrum*) in South Carolina. *Weed Science* 58: 402-407.

NOVAK, S. J., AND R. N. MACK. 2005. Genetic variation in *Bromus tectorum* (Poaceae): comparison between native and introduced populations. *Heredity* 71: 167-176.

OWEN, M. D. K. 2008. Weed species shifts in glyphosate-resistant crops. *Pest Management Science* 64: 377-387.

OWEN, M. D. K., AND I. A. ZELAYA. 2005. Herbicide-resistant crops and weed resistance to herbicides. *Pest Management Science* 61: 301-311.

PARKER, I. M., J. RODRIGUEZ, AND M. E. LOIK. 2003. An evolutionary approach to understanding the biology of invasions: local adaptation and general-purpose genotypes in the weed *Verbascum thapsus*. *Conservation Biology* 17: 1-15.

PIMENTEL, D., R. ZUNIGA, AND D. MORRISON. 2005. Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics* 52: 273-288.

POWLES, S. B., AND Q. YU. 2010. Evolution in action: plants resistant to herbicides. *Annual Review of Plant Biology* 61: 317-347.

REAGON, M., C. S. THURBER, B. L. GROSS, K. M. OLSEN, Y. JIA, AND A. L. CAICEDO. 2010. Genomic patterns of nucleotide diversity in divergent populations of U.S. weedy rice. *BMC Evolutionary Biology* 10: 180.

REFSELL, D. E., AND R. G. HARTZLER. 2009. Effect of tillage on common waterhemp (*Amaranthus rudis*) emergence and vertical distribution of seed in the soil. *Weed Technology* 23: 129-133.

SAKAI, A. K., F. W. ALLENDORF, J. S. HOLT, D. M. LODGE, J. MOLOFSKY, K. A. WITH, S. BAUGHMAN, ET AL. 2001. The population biology of invasive species. *Annual Review of Ecology and Systematics* 32: 305-332.

SAUER, J. D. 1955. Revision of the dioecious amaranths. *Madroño* 13: 5-46.

SAUER, J. D. 1957. Recent migration and evolution of the dioecious amaranths. *Evolution* 11: 11-31.

SCHIERENBECK, K. A., AND N. C. ELLSTRAND. 2009. Hybridization and the evolution of invasiveness in plants and other organisms. *Biological Invasions* 11: 1093-1105.

SCHLAEPFER, D. R., M. GLÄTTLI, M. FISCHER, AND M. VAN KLEUNEN. 2010. A multi-species experiment in their native range indicates pre-adaptation of invasive alien plant species. *New Phytologist* 185: 1087-1099.

SCOTT, J. W., S. E. MEYER, K. R. MERRILL, AND V. J. ANDERSON. 2010. Local population differentiation in *Bromus tectorum* L. in relation to habitat-specific selection regimes. *Evolutionary Ecology* 24: 1061-1080.

SHIVRAIN, V. K., N. R. BURGOS, D. R. GEALY, K. L. SMITH, R. C. SCOTT, A. MAUROMOUSTAKOS, AND H. BLACK. 2009. Red rice (*Oryza sativa*) emergence characteristics and influence on rice yield at different planting dates. *Weed Science* 57: 94-102.

SIBONY, M., AND B. RUBIN. 2002. The ecological fitness of ALS-resistant *Amaranthus retroflexus* and multiple-resistant *Amaranthus blitoides*. *Weed Research* 43: 40-47.

SNOW, A. A., T. M. CULLEY, L. G. CAMPBELL, P. M. SWEENEY, S. G. HEGDE, AND N. C. ELLSTRAND. 2010. Long-term persistence of crop alleles in weedy populations of wild radish (*Raphanus raphanistrum*). *New Phytologist* 186: 537-548.

STECKEL, L. E. 2007. The dioecious *Amaranthus* spp.: here to stay. *Weed Technology* 21: 567-570.

STECKEL, L. E., C. L. SPRAGUE, AND A. G. HAGER. 2002. Common waterhemp (*Amaranthus rudis*) control in corn (*Zea mays*) with single preemergence and sequential applications of residual herbicides. *Weed Technology* 16: 755-761.

STECKEL, L. E., AND C. L. SPRAGUE. 2004. Late-season common waterhemp (*Amaranthus rudis*) interference in narrow- and wide-row soybean. *Weed Technology* 18: 947-952.

SWANTON, C. J., D. R. CLEMENTS, AND D. A. DERKSEN. 1993. Weed succession under conservation tillage: a hierarchical framework for research and management. *Weed Technology* 7: 286-297.

TARDIF, F. J., I. RAJCAN, AND M. COSTEA. 2005. A mutation in the herbicide target site acetohydroxyacid synthase produces morphological and structural alterations and reduces fitness in *Amaranthus powellii*. *New Phytologist* 169: 251-264.

TRANEL, P. J., AND F. TRUCCO. 2009. 21st century weed science: a call for *Amaranthus* genomics. *In* C. N. Stewart Jr. [ed.], Weedy and Invasive Plant Genomics, 53-81. Blackwell Publishing, Ames, IA, USA.

TRANEL, P. J., AND T. R. WRIGHT. 2002. Resistance of weeds to ALS-inhibiting herbicides: what have I learned? *Weed Science* 50: 700-712.

VAN KLEUNEN, M., D. R. SCHLAEPFER, M. GLAETTLI, AND M. FISCHER. 2011. Preadapted for invasiveness: do species traits or their plastic response to shading differ between invasive and non-invasive plant species in their native range? *Journal of Biogeography* 38: 1294-1304.

VIGUEIRA, C. C., K. M. OLSEN, AND A. L. CAICEDO. 2013. The red queen in the corn: agricultural weeds as models of rapid adaptive evolution. *Heredity* 110: 303-311.

WARWICK, S. I., M.-J. SIMARD, A. LÉGÈRE, H. J. BECKIE, L. BRAUN, B. ZHU, P. MASON, ET AL. 2003. Hybridization between transgenic *Brassica napus* L. and its wild relatives: *Brassica rapa* L., *Raphanus raphanistrum* L., *Sinapis arvensis* L., and *Erucastrum gallicum* (Willd.) O. E. Schulz. *Theoretical and Applied Genetics* 107: 528-539.

WILLIAMS, J. L., H. AUGE, AND J. L. MARON. 2008. Different gardens, different results: native and introduced populations exhibit contrasting phenotypes across common gardens. *Oecologia* 157: 239-248.

TABLES AND FIGURES

Table 3.1. Populations of *Amaranthus tuberculatus* used common garden experiments. Region, genetic region, and population numbers correspond to figures 3.2-3.11.

Region of origin (number)	**Genetic region of origin (number)**	**Population name / number (common garden year)**	**State where population is located**	**Population GPS location**
Plains (1)	Western (1)	ADA / 1	OK	N 36.48063, W 95.25453
Plains	Western	CHE / 2	KS	N 37.74475, W 97.78386
Plains	Western	Loup / 3	NE	N 41.41872, W 97.36353
Plains	Western	NE City / 4	NE	N 40.68752, W 95.83635
Plains	Western	Salt River / 5	OK	N 36.77166, W 98.03800
Plains	Western	TCL / 6	KS	N39.43923, W 96.71025
Mississippi Valley (2)	Western	Iowa 1 / 7	IA (USDA PI 603872)	N 42.49167, W 95.97795
Mississippi Valley	Western	Iowa 2 / 8	IA (USDA PI 553086)	N 42.03696, W 93.92773
Mississippi Valley	Eastern (2)	KEY / 9	IL	N 38.73371, W 89.27585
Mississippi Valley	Eastern	KNK / 10	IL	N 41.16075, W 87.62755
Mississippi Valley	Western	MIS / 11	MO	N 38.87902, W 90.18393
Mississippi Valley	Western	SCRR / 12	MO	N 38.78155, W 90.46896
Northeast (3)	Eastern	PTC / 13 (2010)	OH	N 41.51450, W 82.93943
Northeast	Eastern	BTL / 13 (2011)	OH	N 39.42743, W 84.54071
Northeast	Eastern	ION / 14	MI	N 42.97538, W 85.07140
Northeast	Eastern	NEV / 15	OH	N 38.80763, W 84.21171
Northeast	Eastern	PCL / 16	OH	N 39.26801, W 83.38861
Northeast	Eastern	DMD/KMZ / 17 (2010)	MI/MI	N 42.64500, W 84.64970; N 42.44408, W 85.63737
Northeast	Eastern	SCIO / 17 (2011)	OH	N 40.17745, W 83.12640
Northeast	Eastern	York/DEL / 18	ON/ON	N 43.02070, W 79.89105; N 42.93375, W 81.42106

Table 3.2. Results from GLM, nonparametric, and Tukey HSD posthoc analyses to test the effect of region of origin on transplant height, flowering height, mature height, height over time, days to flowering, mature branch number, length of longest branch, and dry above-ground biomass. Significant values at P < or = 0.05 are bold.

GLM	**2010**				**Tukey HSD Posthoc**					
Variable	**Region**				**Mean (95% CI)**			**Multiple Comparisons**		
	df (Hypothesis, Error)	**F**	**Chi-square**	**P**	**Region 1**	**Region 2**	**Region 3**	**P (1 vs. 2)**	**P (2 vs. 3)**	**P (3 vs. 1)**
Transplant Height	2, 15.01	3.538†		0.055	2.055 (2.013, 2.097)	2.054 (2.013, 2.097)	1.888 (1.847, 1.930)	0.999	**<0.001**	**<0.001**
Flowering Height (a)	2 (e)		23.480	**<0.001**	56.457 (51.848, 61.067)	59.408 (53.544, 65.273)	40.180 (33.767, 46.594)	0.653	**<0.001**	**<0.001**
Mature Height Univariate (a)	2, 15.81	12.565		**0.001**	102.566 (95.469, 108.448)	118.943 (112.930, 126.004)	74.882 (64.208, 82.024)	**0.001**	**<0.001**	**<0.001**
Height over Time (b)	4, 518	20.624		**<0.001**	37.814 (35.409, 39.818)	43.429 (41.428, 45.990)	27.151 (23.278, 29.390)	**0.001**	**<0.001**	**<0.001**
Days to Flowering	2 (e)		12.237	**0.002**	69.489 (67.774, 71.203)	68.935 (67.061, 70.809)	62.832 (59.970, 65.693)	0.845	**0.003**	**0.002**
Multivariate Mature Data (c)	6, 512	15.542		**<0.001**	n/a	n/a	n/a	n/a	n/a	**n/a**
Mature Height (a,d)	2, 257	39.844		**<0.001**	104.076 (97.152, 109.845)	120.883 (115.070, 128.205)	74.363 (62.504, 80.341)	**0.001**	**<0.001**	**<0.001**
Mature Branch Number (d)	2, 257	15.565†		**<0.001**	5.077 (4.755, 5.346)	5.804 (5.533, 6.145)	4.602 (4.015, 4.846)	**0.002**	**<0.001**	0.130
Length of Longest Branch (d)	2, 257	0.992†		0.372	n/a	n/a	n/a	n/a	n/a	**n/a**
Mature Branch Number (Univariate)	2, 15.90	6.231†		**0.010**	5.077 (4.755, 5.346)	5.805 (5.533, 6.145)	4.601 (4.015, 4.846)	**0.002**	**<0.001**	0.130
Length of Longest Branch (Univariate)	2, 17.35	0.974†		0.397	n/a	n/a	n/a	n/a	n/a	**n/a**
Dry Above-Ground Biomass	2, 15.84	5.809*		**0.013**	0.786 (0.704, 0.852)	0.981 (0.900, 1.056)	0.711 (0.575, 0.778)	**0.001**	**<0.001**	0.439
GLM	**2011**				**Tukey HSD Posthoc Test**					
Variable	**Region**				**Mean (95% CI)**			**Multiple Comparisons**		
	df (Hypothesis, Error)	**F**	**Chi-square**	**P**	**Region 1**	**Region 2**	**Region 3**	**P (1 vs. 2)**	**P (2 vs. 3)**	**P (3 vs. 1)**
Transplant Height	2, 15.01	3.991		**0.041**	9.428 (9.009, 9.741)	10.004 (9.639, 10.368)	7.795 (7.430, 8.159)	0.073	**<0.001**	**<0.001**
Flowering Height (a)	2, 15.07	6.365		**0.010**	90.146 (85.910, 94.683)	96.025 (91.931, 100.144)	61.597 (54.588, 64.399)	0.129	**<0.001**	**<0.001**
Mature Height Univariate (a)	2, 15.11	6.073		**0.012**	110.495 (105.067, 116.922)	113.107 (107.558, 118.694)	74.212 (64.766, 78.218)	0.800	**<0.001**	**<0.001**
Height over Time (b)	4, 762	31.552		**<0.001**	76.704 (73.752, 80.108)	79.888 (76.950, 82.901)	52.797 (47.560, 54.668)	0.317	**<0.001**	**<0.001**
Days to Flowering	2 (e)		11.542	**0.003**	58.416 (57.240, 59.590)	59.837 (58.720, 60.960)	55.496 (53.510, 57.480)	0.083	**0.001**	**0.043**
Multivariate Mature Data (c)	6, 756	33.882		**<0.001**	n/a	n/a	n/a	n/a	n/a	**n/a**
Mature Height (a,d)	2, 379	52.104		**<0.001**	110.495 (105.044, 116.982)	113.700 (108.120, 119.297)	74.212 (64.696, 78.242)	0.717	**<0.001**	**<0.001**
Mature Branch Number (d)	2, 379	6.442†		**0.002**	5.908 (5.675, 6.123)	6.441 (6.227, 6.647)	6.124 (5.781, 6.290)	**0.002**	0.135	0.408
Length of Longest Branch (d)	2, 379	21.535†		**<0.001**	8.145 (7.806, 8.486)	8.272 (7.947, 8.584)	6.944 (6.325, 7.096)	0.852	**<0.001**	**<0.001**
Mature Branch Number (Univariate)	2 (e)		8.477	**0.014**	36.530 (33.930, 39.120)	43.140 (40.280, 46.010)	40.370 (36.460, 44.270)	**0.002**	0.294	0.231
Length of Longest Branch (Univariate)	2 (e)		17.382	**<0.001**	71.658 (65.591, 77.724)	72.702 (67.196, 78.208)	54.682 (48.276, 61.087)	0.827	**<0.001**	**<0.001**
Dry Above-Ground Biomass	2 (e)		17.516	**<0.001**	56.857 (47.318, 66.397)	58.384 (48.236, 68.532)	35.705 (27.562, 43.848)	0.928	**<0.001**	**<0.001**

*log10 transformed data, †square-root transformed data, a = with transplant height subtracted, b = results for time*region interaction with multivariate Pillai's Trace test, c = results for region with multivariate Pillai's Trace test, d = results from multivariate general linear model, e = could not be transformed to normality, analyzed with Kruskal-Wallis or Mann-Whitney U test.

Table 3.3. Results from GLM and nonparametric analyses to test the effect of sex on flowering height, mature height, height over time, days to flowering, mature branch number, length of longest branch, and dry above-ground biomass. Significant values at P < or = 0.05 are bold.

GLM	**2010**				**2011**			
Variable	**Sex**				**Sex**			
	df (Hypothesis, Error)	**F**	**Z**	**P**	**df (Hypothesis, Error)**	**F**	**Z**	**P**
Flowering Height (a)	1 (e)		-5.772	**<0.001**	1, 380	3.068		0.081
Mature Height Univariate (a)	1, 256	14.816		**<0.001**	1, 377	268.228		**<0.001**
Height over Time (b)	2, 257	30.125		**<0.001**	2, 379	437.173		**<0.001**
Days to Flowering	1 (e)		-5.941	**<0.001**	1 (e)		-4.656	**<0.001**
Multivariate Mature Data (c)	3, 254	42.508		**<0.001**	3, 376	98.634		**<0.001**
Mature Height (a,d)	1, 256	14.816		**<0.001**	1, 378	270.983		**<0.001**
Mature Branch Number (d)	1, 256	100.082†		**<0.001**	1, 378	188.390†		**<0.001**
Length of Longest Branch (d)	1, 256	0.148†		0.700	1, 378	204.079†		**<0.001**
Mature Branch Number (Univariate)	1, 256	100.082†		**<0.001**	1 (e)		-11.166	**<0.001**
Length of Longest Branch (Univariate)	1, 256	0.148†		0.700	1 (e)		-11.725	**<0.001**
Dry Above-Ground Biomass	1, 274	59.649*		**<0.001**	1 (e)		-13.663	**<0.001**

*log transformed data, †square-root transformed data

a = with transplant height subtracted, b = results for time*region (or time*region, or time*genregion) interaction with multivariate Pillai's Trace test, c = results for region with multivariate Pillai's Trace test, d = results from multivariate generallinear model, e = could not be transformed to normality, analyzed with nonparametric tests (Kruskal-Wallis or Mann-Whitney U).

Table 3.4. Results from GLM, nonparametric, and Tukey HSD posthoc analyses to test the effect of region for just female plants on flowering height, mature height, height over time, days to flowering, mature branch number, length of longest branch, and dry above-ground biomass. Significant values at P < or = 0.05 are bold.

GLM	**2010**				**Tukey HSD Posthoc**					
Variable	**Region**				**Mean (95% CI)**		**Multiple Comparisons**			
	df (Hypothesis, Error)	**F**	**Chi-square**	**P**	**Region 1**	**Region 2**	**Region 3**	**P (1 vs. 2)**	**P (2 vs. 3)**	**P (3 vs. 1)**
Flowering Height (a)	2, 15.38	4.450		**0.030**	63.708 (56.587, 68.828)	73.481 (66.762, 79.464)	46.721 (39.659, 52.496)	0.060	**<0.001**	**<0.001**
Mature Height Univariate (a)	2, 16.60	9.613		**0.002**	109.971 (100.112, 117.020)	131.372 (121.029, 138.956)	79.094 (61.917, 85.734)	**0.001**	**<0.001**	**<0.001**
Height over Time (b)	4, 258	11.601		**<0.001**	39.969 (36.556, 42.398)	47.292 (43.713, 49.907)	29.022 (23.097, 31.327)	**0.001**	**<0.001**	**<0.001**
Days to Flowering	2 (e)		6.004	**0.050**	72.319 (70.160, 74.480)	73.394 (71.290, 75.500)	67.050 (63.550, 70.550)	0.255	**0.022**	0.112
Multivariate Mature Data (c)	6, 254	11.226		**<0.001**	n/a	n/a	n/a	n/a	n/a	n/a
Mature Height (a,d)	2, 128	27.373		**<0.001**	109.971 (100.057, 116.972)	130.492 (120.595, 138.602	79.094 (61.946, 85.772)	**0.002**	**<0.001**	**<0.001**
Mature Branch Number (d)	2, 128	11.790		**<0.001**	33.760 (29.682, 37.613)	44.430 (40.657, 49.100)	32.500 (23.877, 35.048)	**0.001**	**0.001**	0.914
Length of Longest Branch (d)	2, 128	0.529†		0.590	n/a	n/a	n/a	n/a	n/a	n/a
Mature Branch Number (Univariate)	2, 17.16	5.570		**0.014**	33.760 (29.694, 37.692)	45.000 (40.984, 49.463)	32.500 (23.800, 35.066)	**<0.001**	**<0.001**	0.915
Length of Longest Branch (Univariate)	2, 23.97	0.511†		0.606	n/a	n/a	n/a	n/a	n/a	n/a
Dry Above-Ground Biomass	2, 17.20	5.000*		**0.019**	0.916 (0.806, 0.997)	1.157 (1.067, 1.277)	0.867 (0.685, 0.964)	**0.002**	**0.001**	0.804
GLM	**2011**				**Tukey HSD Posthoc Test**					
Variable	**Region**				**Mean (95% CI)**		**Multiple Comparisons**			
	df (Hypothesis, Error)	**F**	**Chi-square**	**P**	**Region 1**	**Region 2**	**Region 3**	**P (1 vs. 2)**	**P (2 vs. 3)**	**P (3 vs. 1)**
Flowering Height (a)	2, 17.68	3.939		**0.038**	91.973 (84.284, 97.760)	101.188 (93.902, 107.669)	73.841 (58.520, 80.689)	0.117	**<0.001**	**0.002**
Mature Height Univariate (a)	2, 18.62	8.813		**0.002**	136.937 (126.702, 142.220)	148.630 (140.567, 156.420)	106.600 (88.621, 114.149)	0.075	**<0.001**	**<0.001**
Height over Time (b)	4, 276	8.562		**<0.001**	86.156 (80.716, 89.611)	92.801 (88.196, 97.284)	68.250 (56.834, 71.467)	0.078	**<0.001**	**<0.001**
Days to Flowering	2 (e)		9.295	**0.010**	58.984 (57.160, 60.810)	62.193 (60.730, 63.660)	60.051 (57.070, 63.030)	**0.002**	0.473	0.072
Multivariate Mature Data (c)	6, 274	13.429		**<0.001**	n/a	n/a	n/a	n/a	n/a	n/a
Mature Height (a,d)	2, 138	19.330		**<0.001**	136.937 (126.702, 142.220)	148.630 (140.567, 156.420)	106.600 (88.621, 114.149)	0.075	**<0.001**	**<0.001**
Mature Branch Number (d)	2, 138	12.204		**<0.001**	45.710 (41.143, 48.551)	58.650 (54.036, 61.604)	56.440 (48.653, 60.839)	**<0.001**	0.722	**0.001**
Length of Longest Branch (d)	2, 138	3.494		**0.033**	95.435 (86.507, 101.380)	97.763 (90.889, 106.084)	83.277 (67.090, 91.557)	0.892	**0.035**	0.085
Mature Branch Number (Univariate)	2, 18.36	5.219		**0.016**	45.710 (41.143, 48.551)	58.650 (54.036, 61.604)	56.440 (48.653, 60.839)	**<0.001**	0.722	**0.001**
Length of Longest Branch (Univariate)	2, 20.81	2.358		0.119	n/a	n/a	n/a	n/a	n/a	n/a
Dry Above-Ground Biomass	2, 19.63	4.035†		**0.034**	9.425 (8.599, 10.094)	9.990 (9.315, 10.843)	7.856 (6.079, 8.539)	0.510	**0.001**	**0.019**

*log transformed data, †square-root transformed data, a = with transplant height subtracted, b = results for time*region (or time*region, or time*genregion) interaction with multivariate Pillai's Trace test, c = results for region with multivariate Pillai's Trace test, d = results from multivariate general linear model, e = could not be transformed to normality, analyzed with nonparametric tests (Kruskal-Wallis or Mann-Whitney U).

Table 3.5. Results from GLM and nonparametric analyses to test the effect of genetic region of origin on flowering height, mature height, height over time, days to flowering, mature branch number, length of longest branch, and dry above-ground biomass. Significant values at P < or = 0.05 are bold.

GLM	**2010**				**2011**			
Variable	**Genetic Region**				**Genetic Region**			
	df (Hypothesis, Error)	**F**	**Z**	**P**	**df (Hypothesis, Error)**	**F**	**Z**	**P**
Flowering Height (a)	1e		-3.618	**<0.001**	1, 16.15	7.584		**0.014**
Mature Height Univariate (a)	1, 17.12	6.457		**0.021**	1, 16.22	7.556		**0.014**
Height over Time (b)	2, 258	20.002		**<0.001**	2, 380	51.255		**<0.001**
Days to Flowering	1 (e)		-2.091	**0.037**	1 (e)		-1.955	0.051
Multivariate Mature Data (c)	3, 255	18.935		**<0.001**	3, 377	71.206		**<0.001**
Mature Height (a,d)	1, 257	34.472		**<0.001**	1, 379	71.818		**<0.001**
Mature Branch Number (d)	1, 257	6.808†		**0.010**	1, 379	0.606†		0.437
Length of Longest Branch (d)	1, 257	1.187†		0.277	1, 379	34.819†		**<0.001**
Mature Branch Number (Univariate)	1, 17.87	2.059†		0.169	1 (e)		-1.440	0.150
Length of Longest Branch (Univariate)	1, 23.47	1.172†		0.29	1 (e)		-3.801	**<0.001**
Dry Above-Ground Biomass	1, 17.76	1.650*		0.215	1 (e)		-3.599	**<0.001**

*log transformed data, †square-root transformed data

a = with transplant height subtracted, b = results for time*region (or time*region, or time*genregion) interaction with multivariate Pillai's Trace test, c = results for region with multivariate Pillai's Trace test, d = results from multivariate general linear model, e = could not be transformed to normality, analyzed with nonparametric tests (Kruskal-Wallis or Mann-Whitney U).

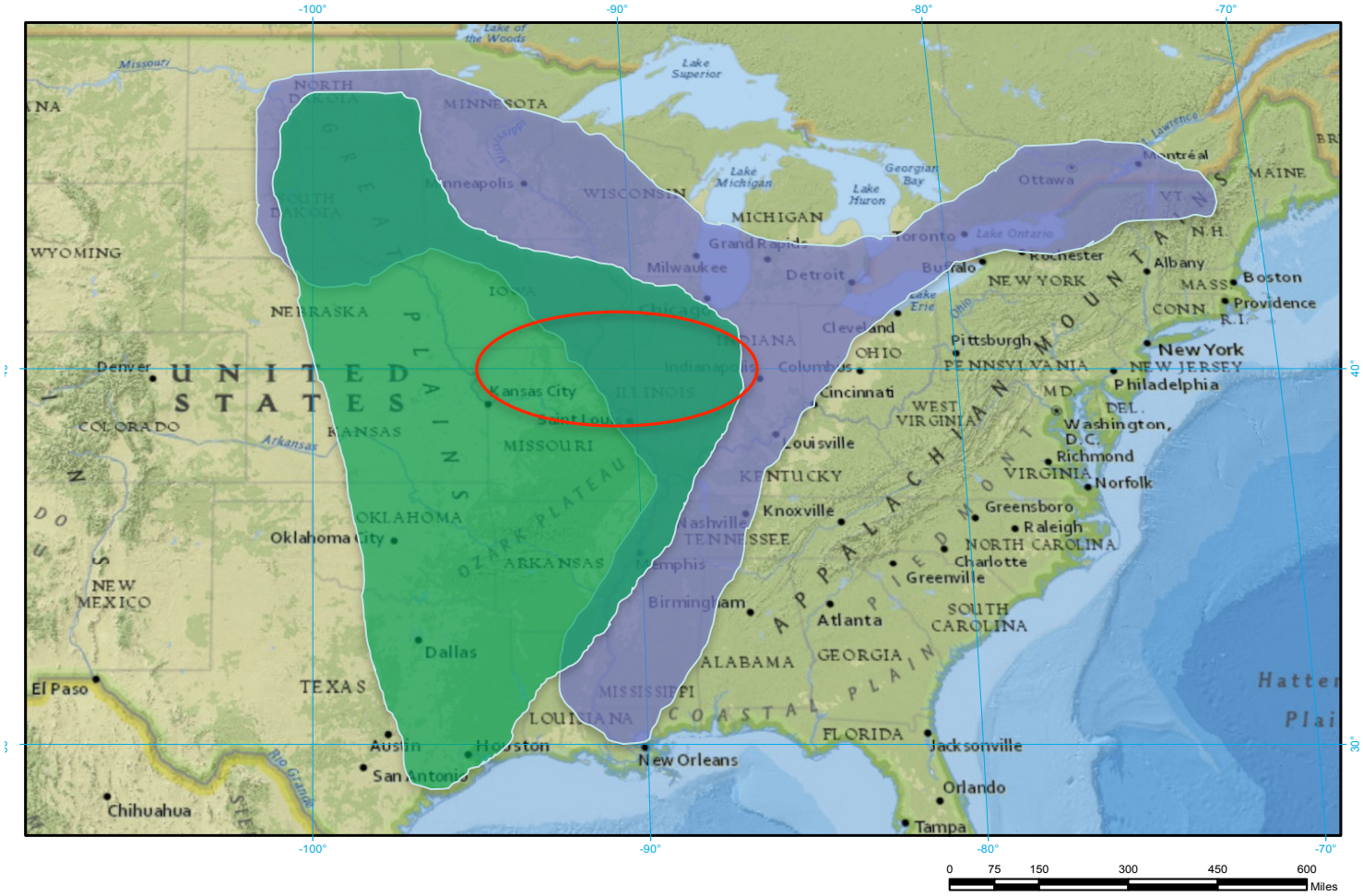

Figure 3.1. Geographical range of *Amaranthus tuberculatus* s.l. (waterhemp), with historical range of *A. tuberculatus* var. *rudis* in green, and range of *A. tuberculatus* var. *tuberculatus* in purple, with the opaque green shading showing the areas of overlap between the varieties (adapted from Sauer, 1957). The red oval surrounds the area of most severe agricultural waterhemp infestation.

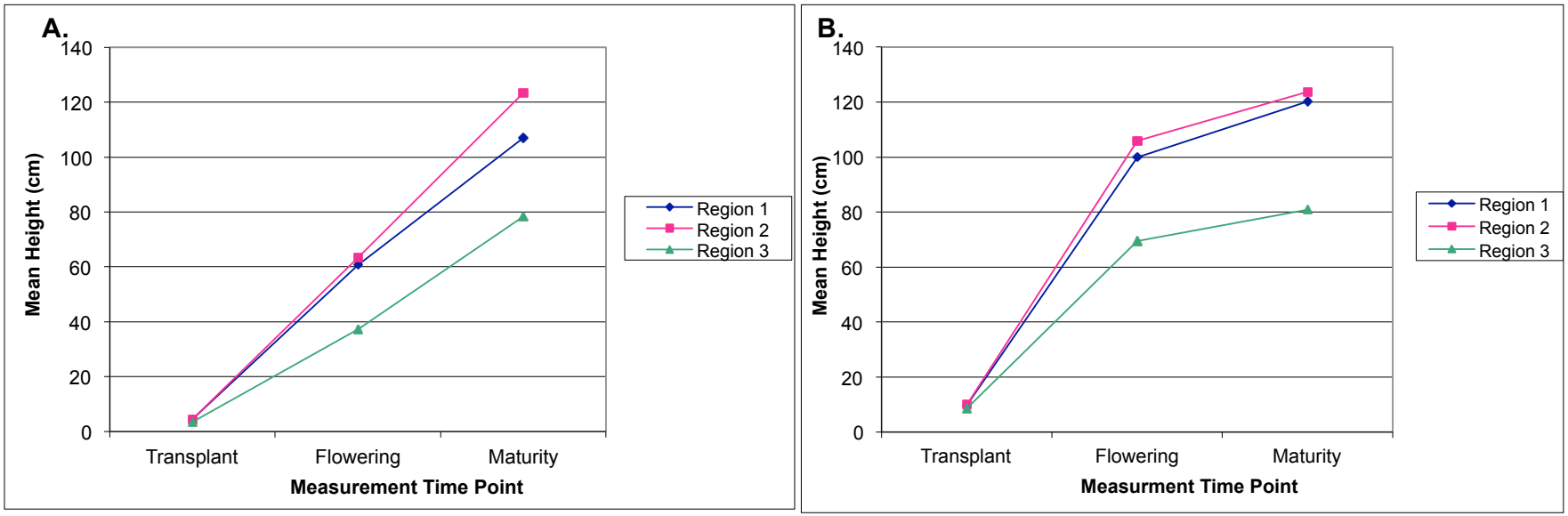

Figure 3.2. Mean height (cm) of plants over time, by region of origin. Region 1 = Plains, Region 2 = Mississippi Valley, Region 3 = Northeast. **A** = 2010, **B** = 2011.

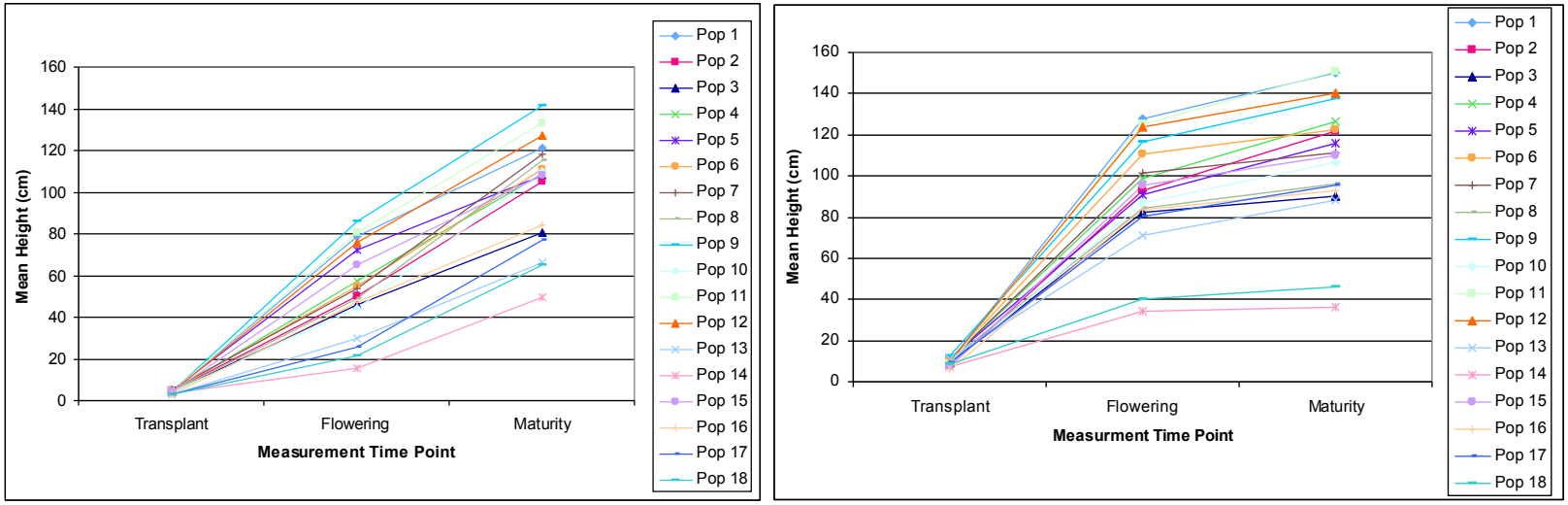

Figure 3.3. Mean height (cm) of plants over time, by population of origin. Region 1 = Pop 1-6, Region 2 = Pop 7-12, Region 3 = Pop 13-18. See Table 3.1 for population names and locations. **A** = 2010, **B** = 2011.

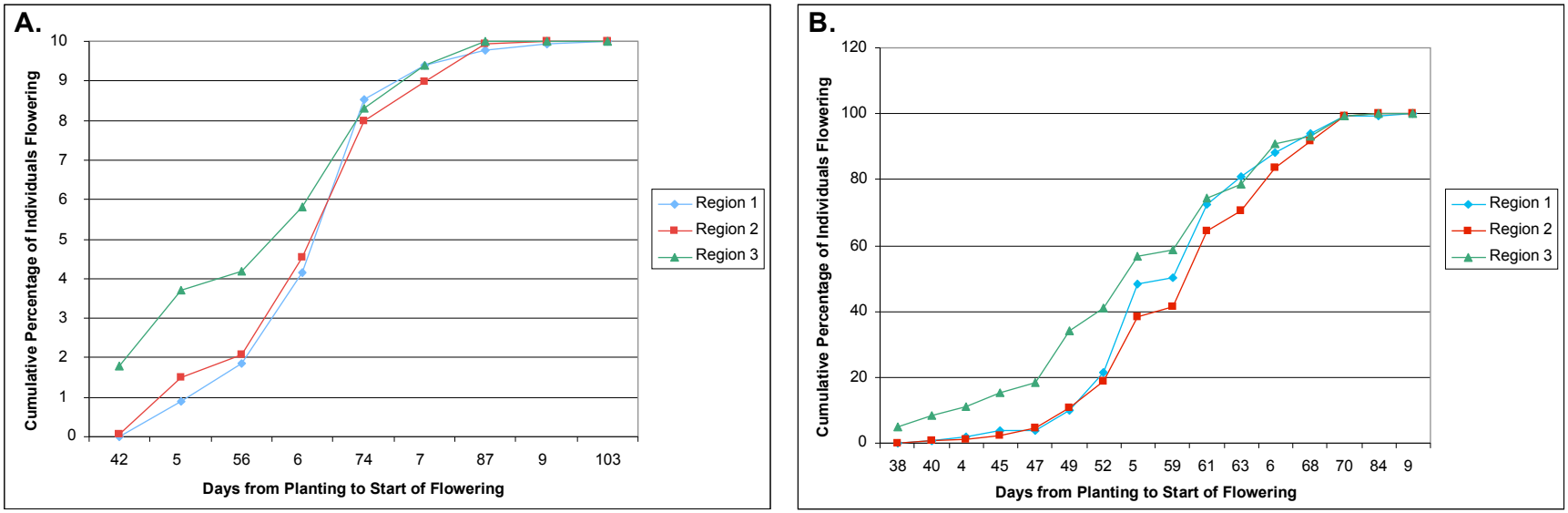

Figure 3.4. Percentage of individuals flowering in each time interval, measured in days from planting to flowering, by region of origin. Region 1 = Plains, Region 2 = Mississippi Valley, Region 3 = Northeast. **A** = 2010, **B** = 2011.

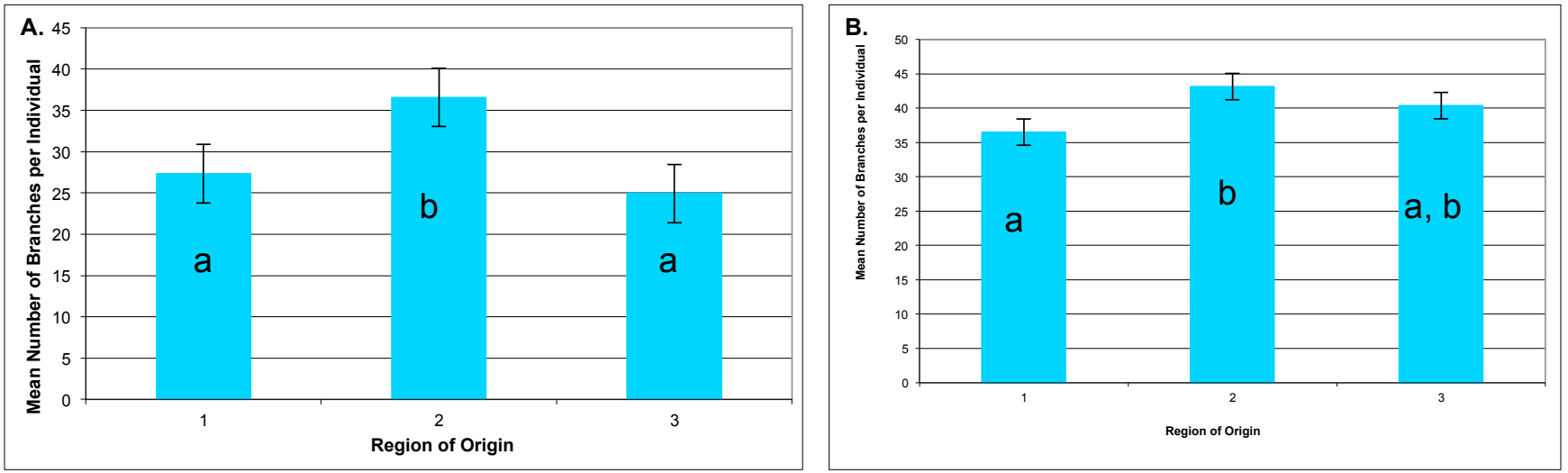

Figure 3.5. Mean number of branches per plant at maturity, by region. Region 1 = Plains, Region 2 = Mississippi Valley, Region 3 = Northeast. Letters next to bars represent groups that are significantly different (different letters) or are not significantly different (same letters) as determined by Tukey HSD tests. **A** = 2010, **B** = 2011.

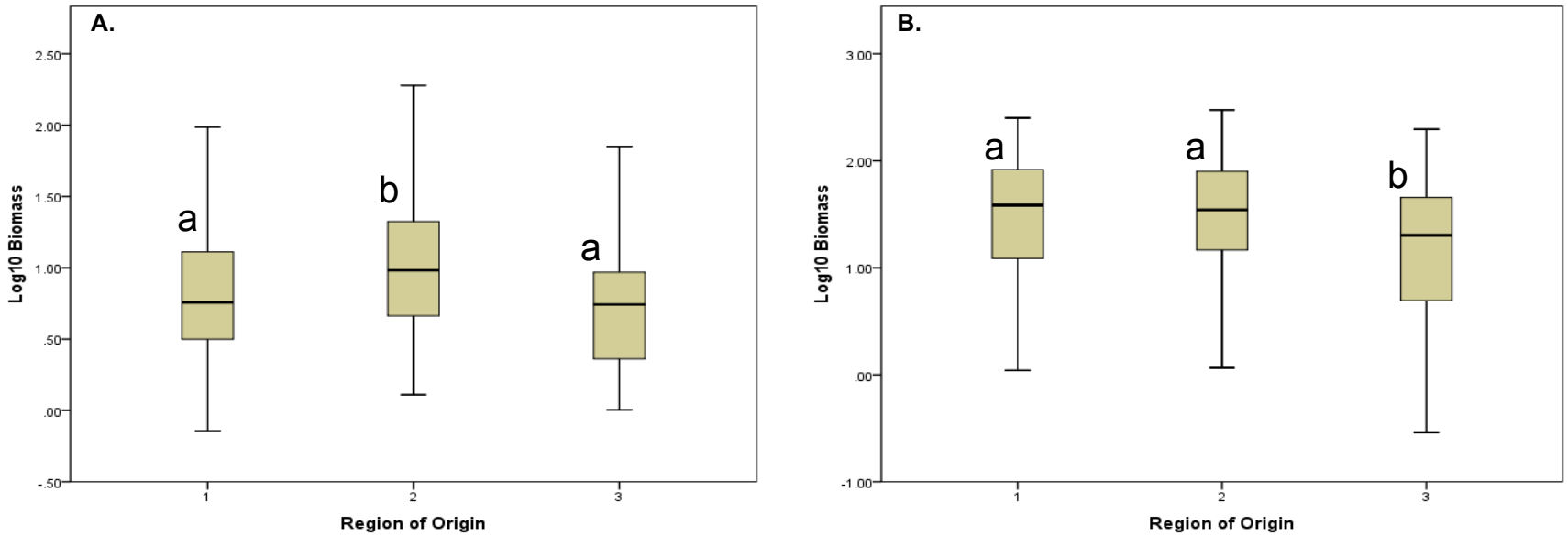

Figure 3.6. Box plots of log10 dry above-ground biomass, by region of origin. Region 1 = Plains, Region 2 = Mississippi Valley, Region 3 = Northeast. Letters on box plots represent groups that are significantly different (different letters) or are not significantly different (same letters) as determined by Tukey HSD tests. **A** = 2010, **B** = 2011.

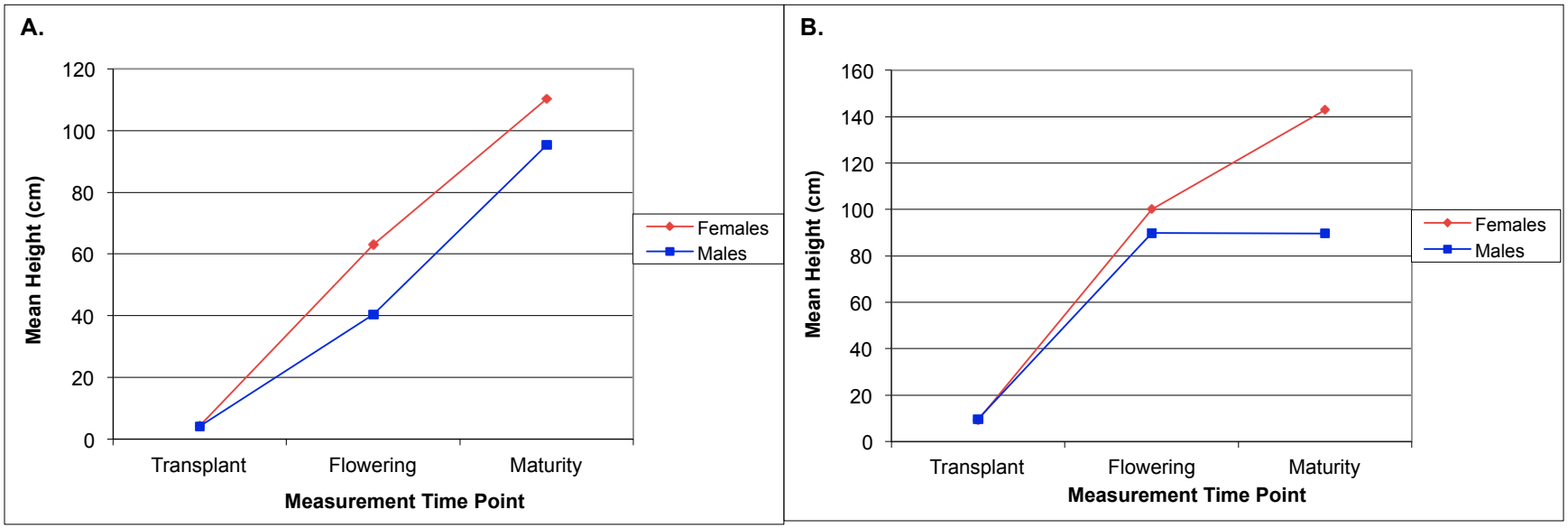

Figure 3.7. Mean height (cm) of plants over time by sex of plants. **A** = 2010, **B** = 2011.

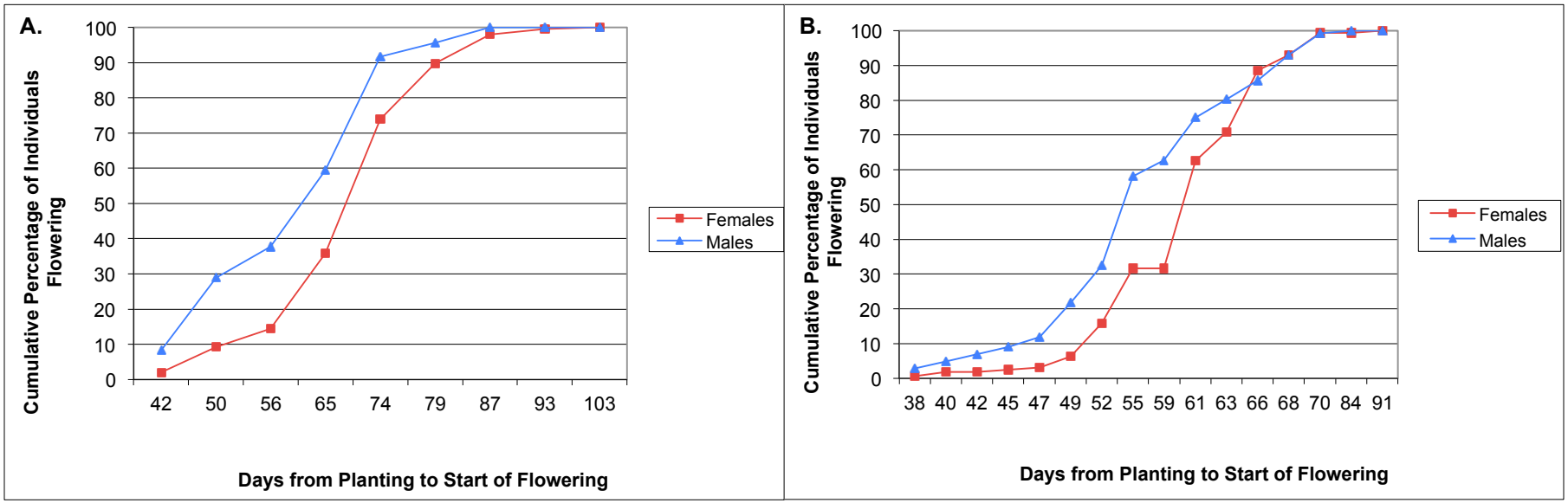

Figure 3.8. Percentage of individuals flowering in each time interval, measured in days from planting to flowering, by sex of plants. **A** = 2010, **B** = 2011.

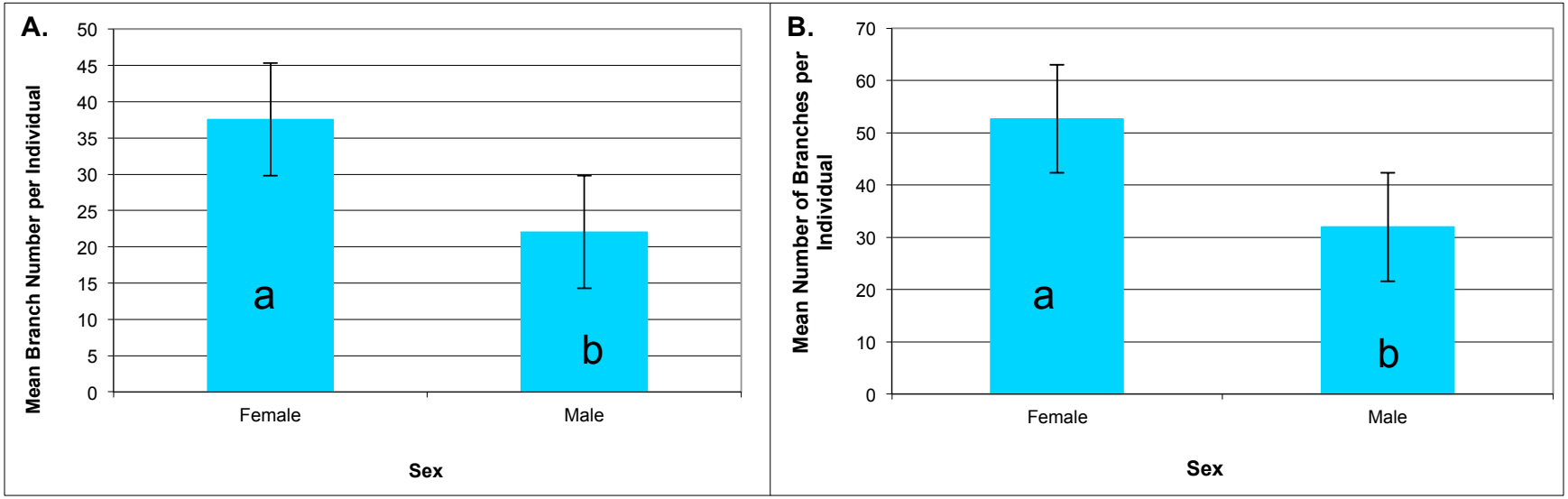

Figure 3.9. Mean number of branches per plant at maturity, by sex of plants. Letters on bars represents groups that are significantly different (different letters) or are not significantly different (same letters). **A** = 2010, **B** = 2011.

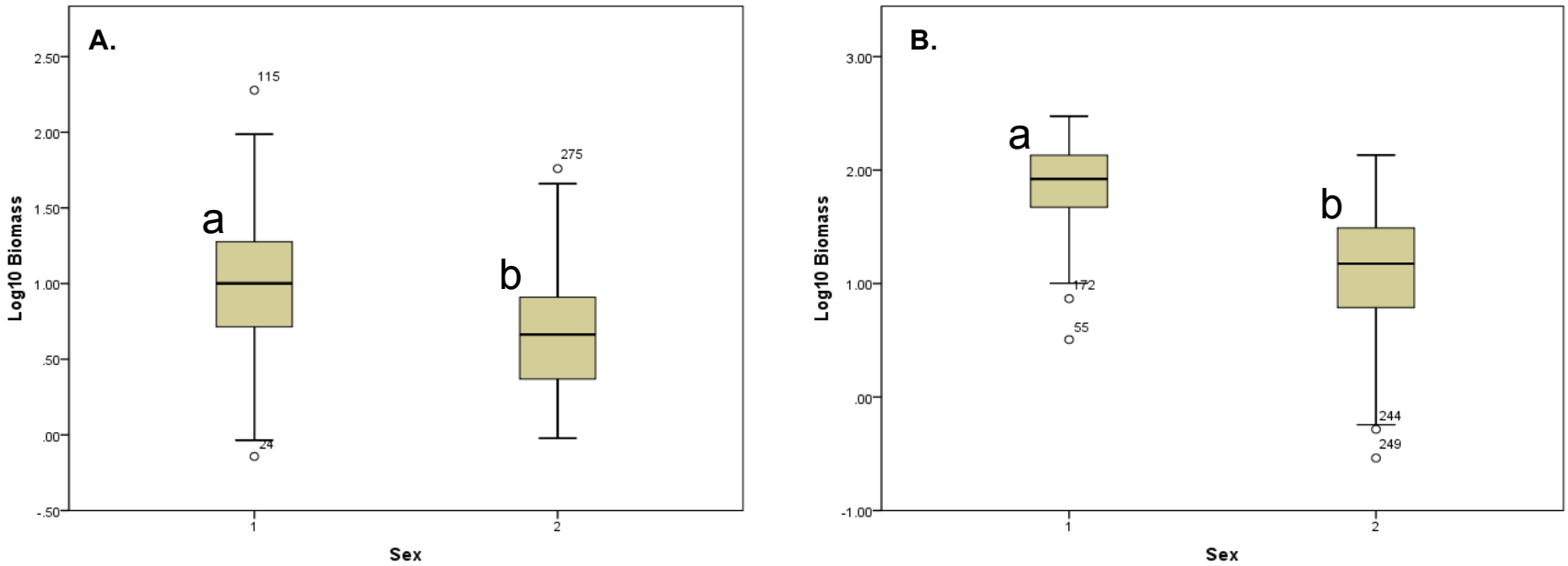

Figure 3.10. Box plots of $\log_{10}$ dry above-ground biomass, by sex of plants. 1 = female, 2 = male. Letters next to box plots represent groups that are significantly different (different letters) or are not significantly different (same letters). **A** = 2010, **B** = 2011.

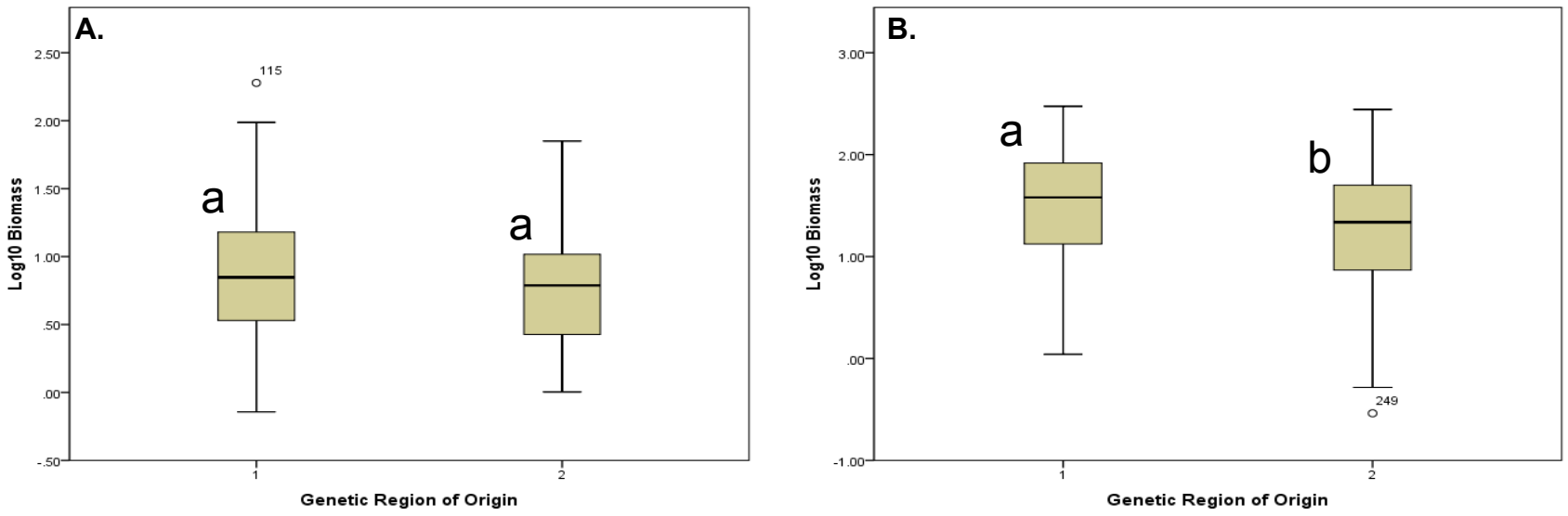

Figure 3.11. Box plots of $\log_{10}$ dry above-ground biomass, by genetic region of origin. Region 1 = "western" subpopulation, Region 2 = "eastern" subpopulation. Letters next to box plots represent groups that are significantly different (different letters) or are not significantly different (same letters). **A** = 2010, **B** = 2011.

CHAPTER 4

Presence of Agriculturally-Adaptive Herbicide Resistance Alleles in Natural Populations of Ohio Waterhemp (*Amaranthus tuberculatus*)

INTRODUCTION

Herbicide resistance in agricultural weeds presents a major economic challenge, with control of herbicide-resistant weeds costing U.S. farmers almost a billion dollars per year (Mortensen, 2010). The population genetics of herbicide resistance evolution has been reviewed and modeled extensively (Maxwell et al., 1990; Warwick, 1991; Jasieniuk and Maxwell, 1994; Jasieniuk et al., 1996), and several factors have been identified that are likely to govern the dynamics of herbicide resistance evolution. These include: the relative fitness of resistant and susceptible genotypes in the absence of herbicides; the intensity of herbicide pressure (which depends on application rates, herbicide rotation practices, and herbicide persistence in the environment); the life history and reproductive system of the weed species; the genetic structure and inheritance of the resistance mutation(s); the frequency of pre-existing herbicide resistance alleles prior to herbicide application; and the potential for augmentation of resistance levels through new mutations and gene flow. Most weed scientists are interested in these evolutionary dynamics out of a desire to control the weed, and thus the studies in this area have been focused on management techniques to slow down the evolution of resistance, such as herbicide and crop rotation and the use of multiple classes of herbicides simultaneously (Jasieniuk and Maxwell, 1994; Diggle et al., 2003).

The role of native, non-agricultural populations of weeds near agricultural weed populations in the evolution of herbicide resistance has seldom been considered. The limited discussion of weeds in natural environments has centered on whether gene flow of susceptible genotypes into a population exposed to herbicide application can slow down the evolution of resistance, with authors disagreeing in their conclusions (Maxwell et al., 1990; Jasieniuk et al., 1996). A related issue, spread of transgenic herbicide resistance by gene flow from crop plants

into their wild or weedy relatives, has also been the subject of concern and research (e.g., in sunflower (Massinga et al., 2005), canola (Snow et al., 1999; Warwick et al., 2008), and rice (Messeguer et al, 2001)). However, the potential for herbicide resistance alleles with little to no fitness cost in natural environments to persist in non-agricultural weed populations, and the implications of this persistence on herbicide resistance evolution in agricultural populations, have never been addressed.

Amaranthus tuberculatus, or waterhemp, the native Midwestern agricultural weed discussed in Chapters 2 and 3, has populations that have evolved resistance to five different chemical classes of herbicides. The following papers reported the first discovery of each type of resistance: Photosystem II inhibitors, also called triazines (Anderson et al., 1996); acetolactate synthase (ALS) inhibitors (Horak and Peterson, 1995); protoporphyrinogen oxidase (PPO) inhibitors (Shoup et al, 2003), p-hydroxyphenylpyruvate dioxygenase (HPPD) inhibitors (Hausman et al., 2011); and glyphosate (Legleiter and Bradley, 2008). Furthermore, some waterhemp populations have developed resistance to multiple herbicide classes (Falk et al., 2005; Patzoldt et al, 2005; McMullan and Green, 2011). The species has extremely large agricultural populations, as it is nearly ubiquitous in agricultural areas of central Midwest and is obligately outcrossing and wind-pollinated (Liu et al., 2012). Waterhemp's constantly evolving resistance patterns make it one of the hardest weeds to control in Midwestern agricultural fields.

One particular class of herbicides, ALS-inhibitors, is not currently recommended to control waterhemp at all, due to the prevalence of resistance throughout the Midwest (Nordby et al., 2010). Acetolactate synthase is an enzyme that catalyzes the synthesis of branched-chain amino acids in all plants; ALS-inhibitors starve the plant of branched-chain amino acids over a short period of time. The first commercial ALS-inhibitor was introduced in 1982, and there are

presently over 50 chemicals in four different ALS-inhibiting subclasses available to farmers (Tranel and Wright 2002). Both pre- and post-emergence chemicals exist in this class, they are effective at very low application rates, and some have soil residual activity (Tranel and Wright, 2002), which is predicted by Jasieniuk and Maxwell (1994) to accelerate the evolution of herbicide resistance.

The molecular basis of ALS resistance in waterhemp has been examined (Foes et al., 1998; Patzoldt and Tranel, 2007), and three single-nucleotide changes to the *ALS* gene have been implicated in resistance. The mutation that causes the broadest spectrum of resistance across different ALS-inhibitor chemicals is a G to T substitution in exon 2 that leads to a Trp574Leu amino acid replacement. This mutation is responsible for much of the ALS-resistance found in natural waterhemp, and is also found in other weed species (Foes et al., 1999; Warwick et al., 2008; Panozzo et al., 2013). The resistant allele is dominant, and there is a perfect correlation between possession of the allele and resistance to ALS-inhibitors (Tranel and Wright, 2002). Two less common ALS resistance mutations in waterhemp occur at amino acid position 653, lead to a serine being replaced by asparagine or threonine, and confer resistance to imidazolinone ALS-inhibitors, but not sulfonyureas (Patzoldt and Tranel, 2007). It is unknown whether any of the three known ALS resistance mutations cause reduced fitness in the absence of herbicide exposure; no fitness trials have been performed in waterhemp, and studies on other *Amaranthus* species have shown variable effects of resistance on fitness (no effects on *A. blitoides* and *A. retroflexus*, Sibony and Rubin, 2002; negative effects on *A. powellii*, Tardif et al., 2006). In plants in general, resistance mutations in the *ALS* gene are not thought to cause consistent reductions in fitness in the absence of herbicide (Holt and Thill, 1994), and one study in lettuce found that resistant individuals had a potential growth rate advantage (Eberlein et al., 1999).

Non-agricultural populations of waterhemp (on riverbanks and lake shores) sometimes occur less than a kilometer from agricultural waterhemp populations in the Midwest (K. Waselkov, pers. obs.). Gene flow between these environments is highly probable, given the potential for long-distance pollen dispersal by wind in this species, and given that *ALS* is a nuclear gene that can be transported by pollen (unlike genes of the maternally inherited plastid genomes); consequently, resistance alleles can be easily spread by pollen-mediated gene flow (Tranel and Wright, 2002; Liu et al., 2012). I was interested in exploring whether "natural" waterhemp populations contained ALS-inhibitor resistance alleles, and if so, at what frequencies the alleles were present in populations at varying distances from crop field waterhemp. The western half of the state of Ohio, which represents the eastern edge of the range of agricultural waterhemp, was chosen as the location to test several hypotheses about herbicide resistance evolution in *A. tuberculatus*. First, I hypothesized that non-agricultural habitats would contain the herbicide resistance mutation, Trp574Leu, which is most prevalent and effective in agricultural waterhemp: this could result either from weak to no selection against the agriculturally-adaptive trait of ALS-inhibitor resistance in these habitats, or from high levels of gene flow from agricultural fields counteracting the effects of strong negative selection in natural habitats.

Second, I hypothesized that the genetic signature of the Ohio invasion of the "weedy" western genetic variety of *A. tuberculatus* and admixture of this variety with the "non-weedy" eastern genetic variety, detectable in microsatellite genotype frequencies, would be reflected in the trait of herbicide resistance as well (see Chapter 2). From this hypothesis, I predicted that Ohio waterhemp populations from inside the "agricultural waterhemp region" would have higher levels of ALS-inhibitor resistance than waterhemp populations outside of the region. Normally,

neutral markers and adaptive traits are expected to show different patterns of population differentiation (McKay and Latta, 2002), but in this case, there is a high probability that ALS resistance has no fitness cost outside crop fields.

MATERIALS AND METHODS

Sample Collection

I collected seeds from 11 populations in the western half of Ohio (Figure 4.1). Collections from three of the populations were made in September 2009, and collections from the remaining eight populations were made in September 2010 (Table 4.1). Seven populations were from the "agricultural waterhemp region" of Ohio, a group of about 10 counties west of Columbus where agricultural fields have been invaded by waterhemp. Of these populations, five were collected from crop fields and two were from natural ecosystems. The remaining Ohio populations were from outside the agricultural waterhemp region: two were located in the southwestern part of the state (100 and 160 km from the agricultural waterhemp region), and two were located in the north/northwestern part (100 and 200 km from the agricultural waterhemp region). Seeds from four parents per field were collected for agricultural waterhemp populations, whereas seeds from 10 parents per field were collected for non-agricultural populations. Table 4.1 shows locality details for each population. The table also shows locality details for two populations from Illinois that are known to be highly resistant or highly susceptible to ALS-inhibitors, which were used as controls in the herbicide screening.

Greenhouse Herbicide Screening

Seeds from each population were screened for ALS-inhibitor resistance with a pre-emergence treatment of imazethapyr. Each population was used in four treatments: two herbicide treatments and two control treatments. For agricultural waterhemp populations (including the two Illinois control populations), 100 seeds from each of the four parents were pooled and divided equally between the four treatments, for 100 seeds/treatment. For non-agricultural waterhemp populations, 40 seeds from each of the 10 parents were pooled and divided equally between the four treatments, for 100 seeds/treatment. For one non-agricultural population (STW), only nine parents were collected, and 80 seeds from the individual STW10 were used in the screening.

Waterhemp shows significant seed dormancy (Leon and Owen, 2003), and must be stratified at low temperatures to mimic winter exposure to ensure good germination in herbicide screening experiments. I used the stratification procedure developed by the Tranel lab at the University of Illinois-UC. Seeds in 1.7 mL Eppendorf tubes were covered in a 1:1 mixture of commercial bleach and water, and soaked for 10 minutes with periodic agitation. The bleach solution was removed and the seeds were rinsed with an equivalent amount of water twice. Then, 0.15% agarose solution was added to cover the seeds, and the tubes were shaken to suspend the seeds. The tubes were subsequently stored at 4ºC for two months, from January 12 to March 11, 2011.

Herbicide screening was performed in collaboration with the Tranel lab, in the greenhouses at the University of Illinois at Urbana-Champaign. On March 11, the seeds were planted in 18-cell flats in a 3:1:1:1 mixture of commercial potting mix (LC1, Sun Gro Horticulture, Canada) to soil to peat to sand. Seeds were placed onto the soil surface with a pipettor, and then sprayed immediately with imazethapyr (Pursuit, BASF) at 1400 g ai/ha (20x

normal field use rate) in a single-dose screen. The flats were all placed in a single room of the greenhouse and watered daily.

Surviving seedlings were counted 10 days after spraying, on March 21. Only living seedlings with cotyledons were counted. From these counts, average percent resistance in the population was estimated by dividing the number of seedlings alive in the herbicide treatment by the number alive in the control treatment, multiplying by 100, and averaging over the two herbicide/control replicates.

A second round of herbicide screening was conducted in 2013 to confirm resistance levels. I picked a subset of five populations that spanned the range of herbicide responses in the first round of screening. Table 4.1 shows which populations were used in both rounds of screening. Seeds were chosen exactly as in the first round, stratified from February 1 to April 1, and planted April 1, 2013. Exactly the same procedure was followed for the herbicide screening, except that seedlings were counted once on April 12 (11 days after planting) and then again on April 15 (14 days after planting) to determine whether herbicide-treated seedlings that appeared stunted relative to those in the control treatment would grow. Slightly different criteria were used to judge resistance: seedlings with at least one true leaf were counted as resistant, rather than seedlings with only cotyledons. The counts from April 15 were used in the percent resistance calculations.

PCR- Restriction Enzyme Assay

On March 24, 2011, plants that had survived the ALS resistance screening were thinned to four individuals per population. On March 28, I collected leaf tissue from these individuals to test with a PCR-restriction enzyme assay for the most common ALS-resistance mutation found

in agricultural waterhemp in Illinois, the Trp574Leu mutation. The populations ABD, CAN, and BTL had no surviving resistant individuals at this time, and the populations GTB, MC, and STW had less than four surviving individuals. Leaf tissue was collected for four individuals each from OTT, PTC, SCIO, DCC, and RT29, two individuals from GTB, and one individual each from STW and MC.

DNA was extracted with Qiagen DNEasy Plant Mini Kits (Qiagen Inc., Valencia, California, USA). A 450-bp section of the ALS gene corresponding to region B of the ALS protein was amplified using the following primers from Foes et al., (1998): ALSF2, 5'-TCCCGGTTAAAATCATGCTC, and ALSR2, 5'-CTAAACGAGAGAACGGCCAG. PCR was performed on ABI GeneAmp 9700 thermocyclers (Applied Biosystems, Carlsbad, California, USA), in 25 uL reactions containing: 1X GoTaq Flexi Buffer (Promega, Madison, Wisconsin, USA), 2 mM MgCl2, 0.2 mM each dNTPs, 0.8 uM each forward and reverse primers, 0.125 uL GoTaq, 13.375 uL nanopure water, and 2 uL genomic DNA. Amplification conditions were: 94ºC for 5 minutes, then 35 cycles of 94 ºC (30 seconds) denaturation, 50 ºC (30 seconds) annealing, 68 ºC (2 minutes) extension, and 72 ºC (7 minutes) final extension.

After amplification, the PCR product was digested with the restriction enzyme MfeI, as described in Foes et al. (1999). This restriction enzyme will cut region B of the ALS gene into two smaller fragments (~30 bp and ~420 bp) if the Trp574Leu resistance mutation is present. The digestion took place in 20 uL reactions containing: 2 uL BSA, 2 uL NEB4, 0.3 uL MfeI, 0.7 uL nanopure water, and 15 uL PCR product. The reactions were incubated at 37C for 2 hours, and then 3 uL undigested PCR product and 3 uL digested product were loaded onto a 2% agarose gel and run at 70V for 1.5 hours to visualize the results.

ALS Sequencing

Finally, gene sequencing was performed on a subset of 13 individuals (with a range of results) to confirm the results of the restriction enzyme assay, and to detect other mutations in region B of ALS. PCR was performed as described above, and PCR cleanup was performed with Invitrogen PureLink Quick PCR Purification Kits (Life Technologies, Carlsbad, California, USA), according to the manufacturer's instructions but starting with 20-25 uL PCR products. Direct sequencing was performed in 12 uL reactions containing: 0.625X sequencing buffer, 0.27 uM primer, 1.0 uL PCR product, 1.0 uL BigDye version 3.0 terminator (Applied Biosystems), and 6.9 uL nanopure water. Sequencing reaction conditions were: 96ºC for 1 minute, then 50 cycles of 96 ºC (10 seconds) denaturation, 50 ºC (5 seconds) annealing, and 60 ºC (4 minutes) extension. Sequences were cleaned with Sephadex columns (GE Healthcare, Piscataway, NJ, USA) and sequenced on the ABI Prism 3130x Genetic Analyzer (Applied Biosystems). All sequences were combined into contigs and quality scores were assigned with the "phred and phrap" function of BioLign 4.0.6.2 (Hall, 2005). After automatic alignment in BioLign, sequence alignments were proofread by eye and edited if necessary.

RESULTS

Greenhouse Herbicide Screening

Results of the herbicide screening (1st and 2nd rounds) are shown in Table 4.2. For round 1, resistance ranged between 0% and 80.7% for Ohio populations, and the sensitive and resistant Illinois control seeds responded as expected, with 0% and 99.9% resistance, respectively. Results of the 1st round of screening are mapped geographically in Figure 4.2, along with whether the population was collected in an agricultural field or a natural habitat. Distance from

the agricultural waterhemp region is not strongly correlated with resistance levels, as the northern Ohio populations had similar resistance levels even though one was twice as far from the region as the other (OTT, 17.8%; PTC, 14.4%), although the closer southern Ohio population did have higher resistance than the farther population (BTL, 33.6%; ABD, 0%). Some populations within the central Ohio agricultural waterhemp region had high levels of resistance, but others had quite low levels (8.1-80.7%). Nor did agricultural or natural populations differ consistently in resistance levels, although agricultural populations had higher resistance on average: the natural populations ranged from 0-63.4% resistance, and the agricultural populations ranged from 21-80.7% resistance. Within the agricultural waterhemp region, where gene flow from crop fields to natural habitats might be expected to be highest, one riverbank population (STW) had quite low levels of resistance (8.1%), while the other natural population (SCIO) had quite high levels (63.4%).

In the 2nd round of screening, replication of the experiment with a subset of the populations yielded similar results for three populations (Table 4.2). The exceptions were the lower resistance observed in the populations GTB (5.1% vs. 60.7% in round 1) and SCIO (34.4% vs. 64.4% in round 1). This is probably because of the different ways that resistant seedlings were counted between the two rounds of screening: seedlings in the cotyledon stage were counted as resistant in the 1st round, whereas for the 2nd round, only seedlings with at least one true leaf were counted. Some GTB and SCIO herbicide-treated seedlings had cotyledons but never grew larger (unlike the seedlings in the control pot), which was the reason for eliminating them in the second round. The lower levels of resistance observed in the second round are thus more likely to be biologically accurate estimates. As with round 1, the sensitive and resistant Illinois control seeds responded as expected.

PCR-Restriction Enzyme Assay

The results of the PCR-restriction genotyping of herbicide-resistant plants are shown in Table 4.3, and a representative gel is shown in Figure 4.3. The digestion produced three types of results: individuals that did not possess the Trp574Leu mutation appeared to have a single band at ~450 bp on the gel. Only a single individual screened did not possess the mutation (GTB 1). Individuals homozygous for the Trp574Leu mutation appeared to have a single band at ~420 bp on the gel (with a second, barely detectable band at 30 bp). Twelve individuals showed this pattern. Finally, individuals heterozygous for the Trp574Leu mutation had one detectable band at ~450 bp, and a second detectable band at ~420 bp (plus a very faint band at 30 bp). Eleven individuals showed this pattern.

Of the genotyped individuals from agricultural waterhemp populations, eight were heterozygous for the Trp574Leu mutation, two were homozygous for the mutation, and one did not possess the mutation. From the non-agricultural populations, 10 individuals were homozygous for the mutation, and three were heterozygous. When the results are examined geographically, the individuals from the agricultural waterhemp region consisted of 10 heterozygotes for the mutation, five homozygotes, and one individual without the mutation. The individuals from outside of this region consisted of one heterozygote for the mutation, and seven homozygotes.

ALS Sequencing

The subset of individuals sequenced showed all three types of results from the restriction enzyme assay. The sequencing confirmed the digest gel results every time. The Trp574Leu

mutation was not present in GTB1, as shown in the digest. Sequencing also identified synonymous mutations in region B of ALS, with seven different mutations present in various individuals, and a single nonsynonymous mutation at base pair 88, in the third position of amino acid 593 of the whole *ALS* gene. This mutation, which changes lysine to asparagine (K to N), was found in almost every individual sequenced, including the single individual that lacked the Trp574Leu mutation. Another study (Patzoldt and Tranel, 2007) detected this mutation in waterhemp and determined that it was not involved in herbicide resistance.

DISCUSSION

I screened 11 Ohio populations of *Amaranthus tuberculatus* from both agricultural and natural habitats to test the hypothesis that agriculturally-adaptive ALS-inhibitor resistance alleles are found in natural habitats. Almost every natural population screened for resistance to ALS-inhibiting herbicides showed some level of resistance, with the exception of the Ohio River population ABD. The primary agricultural waterhemp mutation conferring ALS-inhibitor resistance, Trp574Leu, was found in four of these natural populations. Two of these populations (OTT and PTC) are one hundred and two hundred kilometers (respectively) from the region where agricultural waterhemp is found in Ohio.

The presence of the same allele in natural and agricultural waterhemp populations suggests that gene flow is responsible for its presence in natural habitats, as there are other mutations to the ALS gene that confer some level of ALS-inhibitor resistance (three mutations known in waterhemp, seven known in other *Amaranthus* species (Sibony and Rubin, 2003; Corbett and Tardif, 2008; Tranel et al, 2008)). However, it is possible that the allele arose independently in natural populations and persists at low frequencies as standing variation in the

species, especially if it has no fitness consequences in the absence of ALS-inhibitor application (Sibony and Rubin, 2002; but see Tardif et al., 2006). Although it may be too late to detect constitutive variation in the waterhemp *ALS* gene that existed prior to the widespread use of ALS-inhibitor herbicides, *A. tuberculatus* is known to have standing variation for glyphosate resistance (Zelaya and Owen, 2005; Volenberg et al., 2007). The hypotheses of gene flow vs. *de novo* mutation would be difficult to disentangle in waterhemp, even if the entire *ALS* gene were sequenced for every resistant individual in all 11 Ohio populations, due to the extensive recombination that is believed to occur in waterhemp due to its obligate outcrossing breeding system and large genome size (Rayburn et al., 2005; Thinglum et al., 2011). In a geographic study of PPO-inhibitor resistance in Illinois, Thinglum et al. (2011) found that intragenic recombination appeared to have destroyed any signature of the origins of the resistance alleles.

If the alleles have arisen independently in natural populations, another possibility is that they do not represent neutral, standing variation, but rather are weakly selected for in these habitats. The impact of the Trp574Leu mutation on ALS function is uncertain in waterhemp, given the conflicting results of studies on relative fitness of resistant *Amaranthus* plants (Sibony and Rubin, 2002; Tardif et al., 2006). It is possible that the mutation enhances the performance of plants in the constantly fluctuating environmental conditions of riverbank habitats. More experiments on the fitness consequences of herbicide resistance mutations in the absence of herbicides are needed in most study systems; few have been conducted in waterhemp (but see Duff et al., 2009).

Yet another consideration is that resistance mutations could be present in natural populations because they occasionally or frequently come into contact with herbicides. Some ALS-inhibitors have negative effects on plant fitness at concentrations too low to detect by

standard chemical testing procedures, and bioassays have detected persistence in soil and water for days to years after application (Whitcomb, 1999). Given the close geographical proximity of many riverbank waterhemp populations to crop fields, contact of these plants with agricultural runoff containing ALS-inhibitors is possible. The Trp574Leu mutation is one of the only known ALS mutations that confers resistance to a broad spectrum of ALS-inhibitor chemical subclasses (Tranel and Wright, 2002), making it likely that this specific mutation, if it arose, would be favored preferentially in the presence of repeated exposure to herbicides over multiple years of herbicide rotation in nearby fields. Studies to test for the presence and persistence of herbicides in natural habitats near agricultural fields in Ohio could be illuminating on this point.

On average, the agricultural waterhemp populations had higher ALS-inhibitor resistance than did natural populations in this study, but the levels of resistance reported from the 1st round of screening should be interpreted with caution. The 2nd round of screening probably more accurately captured the actual levels of resistance in each population. The fact that none of the ostensibly herbicide-resistant individuals from the BTL and CAN populations lived long enough to provide useful tissue suggests that these individuals, which never grew past the cotyledon stage, might have had a very weak mechanism of resistance, rather than the Trp574Leu mutation or other effective resistance mutations. When this type of seedling was disregarded in the second round of screening, the levels of resistance were substantially lower than in the first round for several populations.

Even if the agricultural populations DCC, MC, and RT29 have actual resistance levels similar to those reported from the 1st round of screening, two of the agricultural populations still had quite low levels of resistance (GTB and CAN). Furthermore, while one riverbank population (STW) in the agricultural waterhemp region had a very low resistance level, the other

natural population, along the O'Shaughnessy Reservoir near Columbus (SCIO), had a moderate level of resistance (34.4% in the second screening). There were also moderate levels of ALS-inhibitor resistance quite far from the agricultural waterhemp region, in Putnam County (OTT) and Port Clinton (PTC).

The variation in resistance levels among agricultural populations makes sense in light of the recommendations for herbicide-resistant weed control in the Midwest (Hager and Refsell, 2008; Nordby et al., 2010). Due to herbicide rotation practices, which are widely encouraged, levels of resistance to any particular class of herbicides can vary widely from year to year, and differences in rotation schedules among farms can produce spatial variation in resistance as well (Neve et al., 2009; Délye et al., 2010a). Weed scientists also advise against using pure ALS-inhibiting herbicides, such as Classic (DuPont) and Pursuit (BASF), to control waterhemp infestations due to their ineffectiveness (Bradley et al., 2008). Premixes and tank mixes that include ALS-inhibitors as one component are still probably widely used, but many farmers in the Midwest have switched to primarily using glyphosate (Roundup®, Monsanto) due to the popularity of glyphosate-resistant soybeans and corn since their introduction in 1996 and 1998 respectively (Carpenter et al., 2002), which could explain the very low levels of ALS resistance in at least two of my agricultural populations.

From these results, I am unable to distinguish between two alternative explanations for the presence of the Trp574Leu mutation in natural populations. Negative selection against the agriculturally-adaptive trait of ALS-inhibitor resistance is either weak enough that the alleles can persist for long periods of time, or gene flow from agricultural habitats is counteracting stronger selection fairly effectively (Slatkin, 1987). Experimental tests on the relative fitness of waterhemp with resistant and sensitive ALS alleles in herbicide-free environments would be

required to support one alternative or the other. Returning to the question that drove the study, what is the role of natural habitats in herbicide resistance? It appears from my findings as though natural populations could be acting as overlooked genetic "reservoirs" for herbicide resistance alleles. A study of herbicide resistant blackgrass (*Alopecurus myosuroides*) in France revealed that the enormous amount of pollen exchange between neighboring farms for this wind-pollinated weed mean that even if individual farmers successfully eradicated blackgrass from their particular fields by rotating herbicides, they would inevitably "catch" herbicide resistance from nearby, less well-managed agricultural fields (Délye et al., 2010b). In waterhemp, it seems just as likely that herbicide resistant gene flow is coming from nearby riverbanks and lake shores.

Finally, there is a remarkably close correspondence between the findings of Chapter 2 and the levels of herbicide resistance across the state observed in this chapter. Microsatellite marker analyses with TESS and BAPS show admixture between the invading agricultural waterhemp in Ohio and the native waterhemp that naturally occurs on riverbanks, especially within the "agricultural waterhemp region." The STRUCTURE analysis of Ohio populations alone also shows some admixture in the populations to the north and south of this region, except along the Ohio River. The results from the current chapter are in agreement with all of these results, although it is unclear whether the same forces are driving the neutral marker and resistance patterns in the agricultural populations, as variable temporal and spatial herbicide pressure could have led to a parallel pattern in herbicide resistance. The correspondence of the results for the northern and southern populations, however, makes a case for gene flow of the ALS-inhibitor resistance alleles rather than independent evolution within natural populations. Note also that genes that are selected upon in both agricultural and natural ecosystems might

show different patterns from those observed for herbicide resistance: alleles for this trait may be functional only in agricultural habitats.

This study serves as a reminder that management actions in agricultural ecosystems can have effects that spill over into natural ecosystems, and those in turn can reciprocally affect agriculture. Due to the widespread planting of Roundup Ready® crops (92% of U.S. soybean acres in 2010 (Mortensen, 2010)), the overuse or exclusive use of glyphosate is accelerating evolution of resistance to this chemical class in multiple weeds, including waterhemp. Misuse of glyphosate, which is less environmentally persistent and toxic than many older herbicides, is likely to force farmers to return to these more harmful chemicals. For species like waterhemp, in which agricultural populations frequently exchange genes with natural populations that contain alleles resistant to these older herbicides, even this option may not lead to effective weed control.

LITERATURE CITED

ANDERSON, D. D., F. W. ROETH, AND A. R. MARTIN. 1996. Occurrence and control of triazine-resistant common waterhemp (*Amaranthus rudis*) in field corn (*Zea mays*). *Weed Technology* 10: 570-575.

BRADLEY, K., R. SMEDA, AND R. MASSEY. 2008. Management of glyphosate-resistant waterhemp in corn and soybean. Integrated Pest Management MU Guide IPM1030, University of Missouri Extension, Columbia, Missouri, USA. http://extension.missouri.edu/p/IPM1030

CARPENTER, J., A. FELSOT, T. GOODE, M. HAMMIG, D. ONSTAD, AND S. SANKULA. 2002. *Comparative Environmental Impacts of Biotechnology-derived and Traditional Soybean, Corn, and Cotton Crops.* Council for Agricultural Science and Technology, Ames, Iowa, USA. www.cast-science.org. Sponsored by the United Soybean Board. www.unitedsoybean.org

CORBETT, C.-A. L., AND F. J. TARDIF. 2008. Detection of resistance to acetohydroxyacid synthase inhibitors in Amaranthus sp. using DNA polymorphisms. *Pesticide Biochemistry and Physiology* 92: 48-55.

DÉLYE, C., S. MICHEL, A. BÉRARD, B. CHAUVEL, D. BRUNEL, J-P. GUILLEMIN, F. DESSAINT, AND V. LE CORRE. 2010a. Geographical variation in resistance to acetyl-coenzyme A carboxylase-inhibiting herbicides across the range of the arable weed *Alopecurus myosuroides* (black-grass). *New Phytologist* 186: 1005-1017.

DÉLYE, C., J. A. J. CLÉMENT, F. PERNIN, B. CHAUVEL, AND V. LE CORRE. 2010b. High gene flow promotes the genetic homogeneity of arable weed populations at the landscape level. *Basic and Applied Ecology* 11: 504-512.

DIGGLE, A. J., P. B. NEVE, AND F. P. SMITH. 2003. Herbicides used in combination can reduce the probability of herbicide resistance in finite weed populations. *Weed Research* 43: 371-382.

DUFF, M. G., K. AL-KHATIB, AND D. E. PETERSON. 2009. Relative competitiveness of protoporphyrinogen oxidase-resistant common waterhemp (*Amaranthus rudis*). *Weed Science* 57: 169-174.

EBERLEIN, C. V., M. J. GUTTIERI, P. H. BERGER, J. K. FELLMAN, C. A. MALLORY-SMITH, D. C. THILL, R. J. BAERG, AND W. R. BELKNAP. 1999. Physiological consequences of mutation for ALS-inhibitor resistance. *Weed Science* 47: 383-392.

FALK, J. S., D. E. SHOUP, K. AL-KHATIB, AND D. E. PETERSON. 2005. Survey of common waterhemp (*Amaranthus rudis*) response to protox- and ALS-inhibiting herbicides in northeast Kansas. *Weed Technology* 19: 838-846.

FOES, M. J., P. J. TRANEL, L. M. WAX, AND E. W. STOLLER. 1998. A biotype of common waterhemp (*Amaranthus rudis*) resistant to triazine and ALS herbicides. *Weed Science* 46: 514-520.

FOES, M. J., L. LIU, G. VIGUE, E. W. STOLLER, L. M. WAX, AND P. J. TRANEL. 1999. A kochia (*Kochia scoparia*) biotype resistant to triazine and ALS-inhibiting herbicides. *Weed Science* 47: 20-27.

HAGER, A. G., AND D. REFSELL. 2008. Chapter 12: Weed resistance to herbicides. *In* S. Bissonnette [ed.], 2008 Illinois Agricultural Pest Management Handbook, University of Illinois Extension, Urbana-Champaign, Illinois, USA.

HAUSMAN, N. E., S. SINGH, P. J. TRANEL, D. E. RIECHERS, S. S. KAUNDUN, N. D. POLGE, D. A. THOMAS, AND A. G. HAGER. 2011. Resistance to HPPD-inhibiting herbicides in a population of waterhemp (*Amaranthus tuberculatus*) from Illinois, United States. *Pest Management Science* 67: 258-261.

HOLT, J. S. AND D. C. THILL. 1994. Growth and productivity of resistant plants. *In* S. B. Powles and J. A. M. Holtum [eds.], Herbicide Resistance in Plants: Biology and Biochemistry. Lewis Publishers, Boca Raton, Florida, USA, pgs 299–316.

HORAK, M. J., AND D. E. PETERSON. 1995. Biotypes of Palmer amaranth (*Amaranthus palmeri*) and common waterhemp (*Amaranthus rudis*) are resistant to imazethapyr and thifensulfuron. *Weed Technology* 9: 192-195.

JASIENIUK, M., AND B. D. MAXWELL. 1994. Population genetics and the evolution of herbicide resistance in weeds. *Phytoprotection* 75: 25-35.

JASIENIUK, M., A. L. BRÛLÉ-BABEL, AND I. N. MORRISON. 1996. The evolution and genetics of herbicide resistance in weeds. *Weed Science* 44: 176-193.

LEGLEITER, T. R., AND K. W. BRADLEY. 2008. Glyphosate and multiple herbicide resistance in common waterhemp (*Amaranthus rudis*) populations from Missouri. *Weed Science* 56: 582-587.

LEON, R. G., AND M. D. K. OWEN. 2003. Regulation of weed seed dormancy through light and temperature interactions. *Weed Science* 51: 752-758.

LIU, J., A. S. DAVIS, AND P. J. TRANEL. 2012. Pollen biology and dispersal dynamics in waterhemp (*Amaranthus tuberculatus*). *Weed Science* 60: 416-422.

MASSINGA, R. A., K. AL-KHATIB, P. ST. AMAND, AND J. F. MILLER. 2005. Relative fitness of imazamox-resistant common sunflower and prairie sunflower. *Weed Science* 53: 166-174.

MAXWELL, B. D., M. L. ROUSH, AND S. R. RADOSEVICH. 1990. Predicting the evolution and dynamics of herbicide resistance in weed populations. *Weed Technology* 4: 2-13.

MCKAY, J. K., AND R. G. LATTA. 2002. Adaptive population divergence: markers, QTL and traits. *Trends in Ecology and Evolution* 17: 285-291.

MCMULLAN, P. M., AND J. M. GREEN. 2011. Identification of a tall waterhemp (*Amaranthus tuberculatus*) biotype resistant to HPPD-inhibiting herbicides, atrazine, and thifensulfuron in Iowa. *Weed Technology* 25: 514-518.

MESSEGUER, J., C. FOGHER, E. GUIDERDONI, V. MARFÀ, M. M. CATALÀ, G. BALDI, AND E. MELÉ. 2001. Field assessments of gene flow from transgenic to cultivated rice (*Oryza sativa* L.) using a herbicide resistance gene as tracer marker. *Theoretical and Applied Genetics* 103: 1151-1159.

MORTENSEN, D. A. 2010. Statement prepared for the Domestic Policy Subcommittee of the Oversight and Government Reform Committee. House of Representatives Hearing, 111 Congress, Second Session: Are superweeds an outgrowth of USDA biotech policy? Available online: http://oversight.house.gov/wpcontent/uploads/2012/01/20100728Mortensen.pdf

NEVE, P., M. VILA-AIUB, AND F. ROUX. 2009. Evolutionary-thinking in agricultural weed management. *New Phytologist* 184: 783-793.

NORDBY, D., R. HARTZLER, AND K. BRADLEY. 2010. Biology and management of waterhemp. The Glyphosate, Weeds, and Crops Series, by the Glyphosate, Weeds, and Crops Group (http://www.glyphosateweedscrops.org/). Published by Purdue Extension, Purdue University, Lafayette, Indiana, USA.

PANOZZO, S., L. SCARABEL, P. J. TRANEL, AND M. SATTIN. 2013. Target-site resistance to ALS inhibitors in the polyploidy species *Echinochloa crus-galli*. *Pesticide Biochemistry and Physiology* 105: 93-101.

PATZOLDT, W. L., AND P. J. TRANEL. 2007. Multiple ALS mutations confer herbicide resistance in waterhemp (*Amaranthus tuberculatus*). *Weed Science* 55: 421-428.

PATZOLDT, W. L., P. J. TRANEL, AND A. G. HAGER. 2005. A waterhemp (*Amaranthus tuberculatus*) biotype with multiple resistance across three herbicide sites of action. *Weed Science* 53: 30-36.

RAYBURN, A. L., R. MCCLOSKEY, T. C. TATUM, G. A. BOLLERO, M. R. JESCHKE, AND P. J. TRANEL. 2005. Genome size analysis of weedy *Amaranthus* species. *Crop Science* 45: 2557-2562.

SIBONY, M., AND B. RUBIN. 2002. The ecological fitness of ALS-resistant *Amaranthus retroflexus* and multiple-resistant *Amaranthus blitoides*. *Weed Research* 43: 40-47.

SIBONY, M., AND B. RUBIN. 2003. Molecular basis for multiple resistance to acetolactate synthase-inhibiting herbicides and atrazine in *Amaranthus blitoides* (prostrate pigweed). *Planta* 216: 1022-1027.

SHOUP, D. E., K. AL-KHATIB, AND D. E. PETERSON. 2003. Common waterhemp (*Amaranthus rudis)* resistance to protoporphyrinogen oxidase-inhibiting herbicides. *Weed Science* 51: 145-150.

SLATKIN, M. 1987. Gene flow and the geographic structure of natural populations. *Science* 236: 787-792.

SNOW, A. A., B. ANDERSEN, AND R. BAGGER JØRGENSEN. 1999. Costs of transgenic herbicide resistance introgressed from *Brassica napus* into weedy *B. rapa*. *Molecular Ecology* 8: 605-615.

TARDIF, F. J., I. RAJCAN, AND M. COSTEA. 2006. A mutation in the herbicide target site acetohydroxyacid synthase produces morphological and structural alterations and reduces fitness in *Amaranthus powellii*. *New Phytologist* 169: 251-264.

THINGLUM, K. A., C. W. RIGGINS, A. S. DAVIS, K. W. BRADLEY, K. AL-KHATIB, AND P. J. TRANEL. 2011. Wide distribution of the waterhemp (*Amaranthus tuberculatus*) ΔG210 *PPX2* mutation, which confers resistance to PPO-inhibiting herbicides. *Weed Science* 59: 22-27.

TRANEL, P. J., AND T. R. WRIGHT. 2002. Resistance of weeds to ALS-inhibiting herbicides: what have we learned? *Weed Science* 50: 700-712.

TRANEL, P. J., T. R. WRIGHT, AND I. M. HEAP. 2008. Muations in herbicide-resistant weeds to ALS inhibitors. Online: http://www.weedscience.org/Mutations/MutationDisplayAll.aspx. Accessed 6/8/2013.

VOLENBERG, D. S., W. L. PATZOLDT, A. G. HAGER, AND P. J. TRANEL. 2007. Responses of contemporary and historical waterhemp (*Amaranthus tuberculatus*) accessions to glyphosate. *Weed Science* 55: 327-333.

WARWICK, S. I. 1991. Herbicide resistance in weedy plants: physiology and population biology. *Annual Review of Ecology and Systematics* 22: 95-114.

WARWICK, S. I., A. LÉGÈRE, M.-J. SIMARD, AND T. JAMES. 2008. Do escaped transgenes persist in nature? The case of an herbicide resistance transgene in a weedy *Brassica rapa* population. *Molecular Ecology* 17: 1387-1395.

WHITCOMB, C. E. 1999. An introduction to ALS-inhibiting herbicides. *Toxicology and Industrial Health* 15: 232-240.

ZELAYA, I. A., AND M. D. K. OWEN. 2005. Differential response of *Amaranthus tuberculatus* (Moq ex DC) JD Sauer to glyphosate. *Pest Management Science* 61: 936-950.

TABLES AND FIGURES

Table 4.1. Populations of *Amaranthus tuberculatus* from Ohio used in herbicide screening experiments, plus "control" populations from Illinois (chosen because previous experiments demonstrated extremely high or low levels of ALS resistance in these seed stocks).

Population abbreviation	County	Year seeds collected	Agricultural field or riverbank population?	Agricultural waterhemp region?	Used in 2nd round of screening?	Population GPS location
STW	Miami	2010	Riverbank	Yes	Yes	N 40.12163, W 84.35867
GTB	Miami	2010	Soy field	Yes	Yes	N 40.12010, W 84.39868
DCC	Darke	2010	Soy field	Yes	No	N 40.15865, W 84.67002
RT29	Mercer	2010	Corn field	Yes	No	N 40.54592, W 84.63413
MC	Union	2010	Soy field	Yes	No	N 40.15558, W 83.45533
CAN	Madison	2010	Soy field	Yes	No	N 39.98585, W 83.33963
SCIO	Delaware	2010	Lake shore	No	Yes	N 40.17745, W 83.12640
BTL	Butler	2010	Riverbank	No	No	N 39.42743, W 84.54071
ABD	Brown	2009	Riverbank	No	Yes	N 38.65430, W 83.76233
OTT	Putnam	2009	Riverbank	No	No	N 41.03783, W 83.81350
PTC	Ottawa	2009	Lake shore	No	Yes	N 41.51450, W 82.93943
WCS	Wayne Co., IL	1998	Agricultural field	-	Yes	Unknown
ACR	Adams Co., IL	2001	Agricultural field	-	Yes	Unknown

Table 4.2. Results of screening seeds of *Amaranthus tuberculatus* populations for ALS-inhibitor resistance with a single dose of imazethapyr (Pursuit, BASF). Resistance was calculated by dividing # of surviving herbicide-treated seedlings by # of control seedlings and multiplying by 100 for each replicate, and averaging the two replicates. Surviving seedlings were counted slightly differently between the two screening rounds: all green living plants were counted after 11 days in round 1, whereas only living plants with at least one true leaf were counted after 13 days in round 2.

Population abbreviation	**First round (March 2011)**					**Second round (April 2013)**				
	# seedlings alive in herbicide pot 1	**# seedlings alive in herbicide pot 2**	**# seedlings alive in control pot 1**	**# seedlings alive in control pot 2**	**Average % resistance**	**# seedlings alive in herbicide pot 1**	**# seedlings alive in herbicide pot 2**	**# seedlings alive in control pot 1**	**# seedlings alive in control pot 2**	**Average % resistance**
STW	3	10	73	83	8.1	5	3	86	82	4.7
GTB	45	32	69	57	60.7	3	4	66	71	5.1
DCC	32	41	78	70	49.8					
RT29	77	53	81	80	80.7					
MC	33	27	57	62	50.7					
CAN	17	18	82	85	21.0					
SCIO	34	46	62	64	63.4	16	31	76	65	34.4
BTL	20	20	76	49	33.6					
ABD	0	0	75	60	0.0	0	0	68	69	0.0
OTT	17	13	86	82	17.8					
PTC	11	13	82	84	14.4	15	20	59	56	30.6
WCS	0	0	93	83	0.0	0	0	68	100	0.0
ACR	86	74	85	75	99.9	91	94	86	85	100.0

Table 4.3. Results of the restriction enzyme digests of PCR-amplified ALS genes, by population. Individuals are a subset of the individuals in each population that demonstrated resistance during the March 2011 herbicide screening with Pursuit.

Population abbreviation	**Number of heterozygotes for mutation**	**Number of homozygotes for mutation**	**Number of individuals without mutation**	**Total number of individuals screened**
STW	1	0	0	1
GTB	1	0	1	2
DCC	3	1	0	4
RT29	4	0	0	4
MC	0	1	0	1
SCIO	1	3	0	4
OTT	0	4	0	4
PTC	1	3	0	4
Total	11	12	1	24

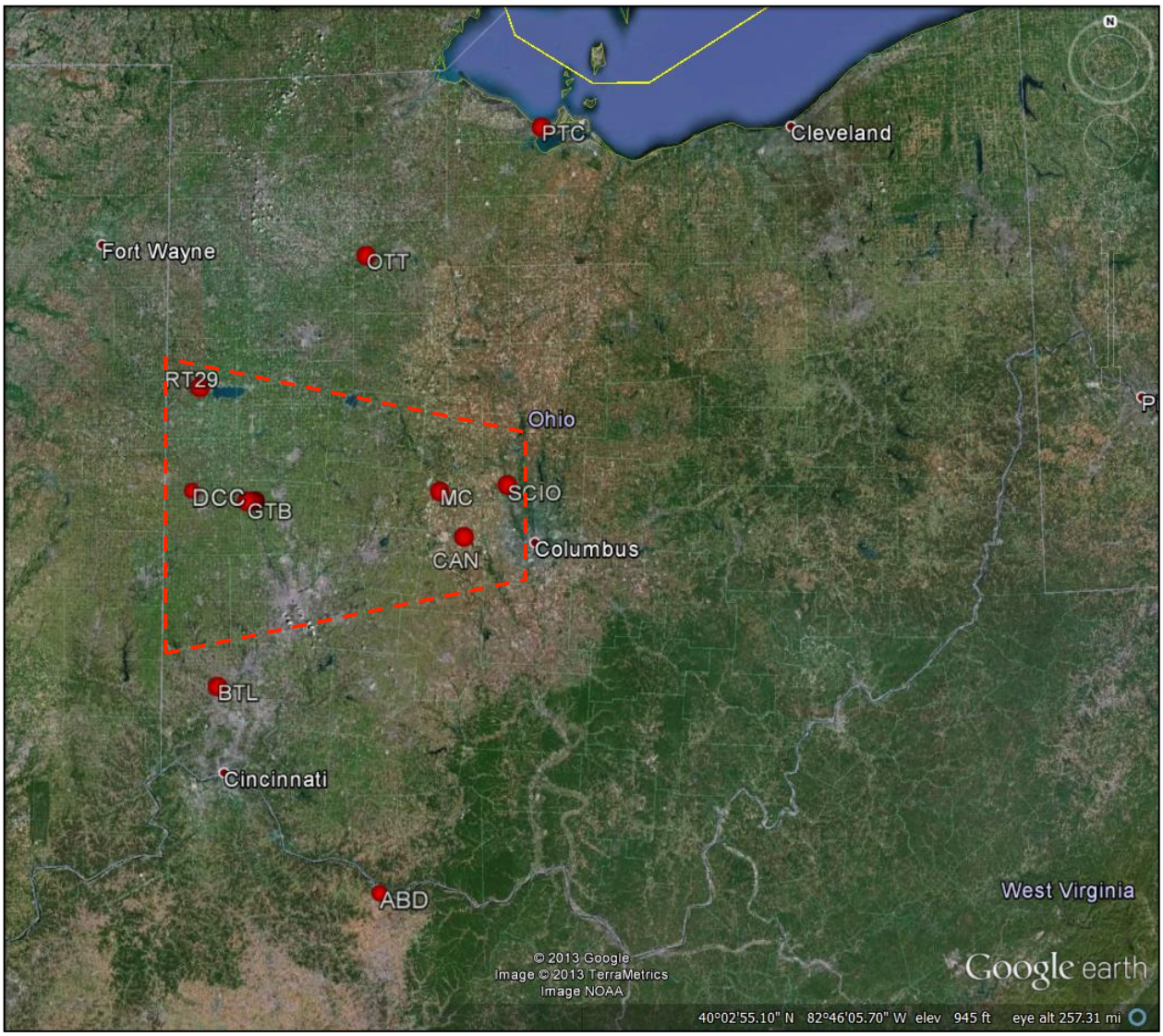

Figure 4.1. Geographical locations of 11 herbicide-screened populations of *Amaranthus tuberculatus* from Ohio. Geographic coordinates were plotted in Google Earth. The population STW is hidden behind GTB on the map. The red dashed trapezoid outlines the "agricultural waterhemp region" of the state.

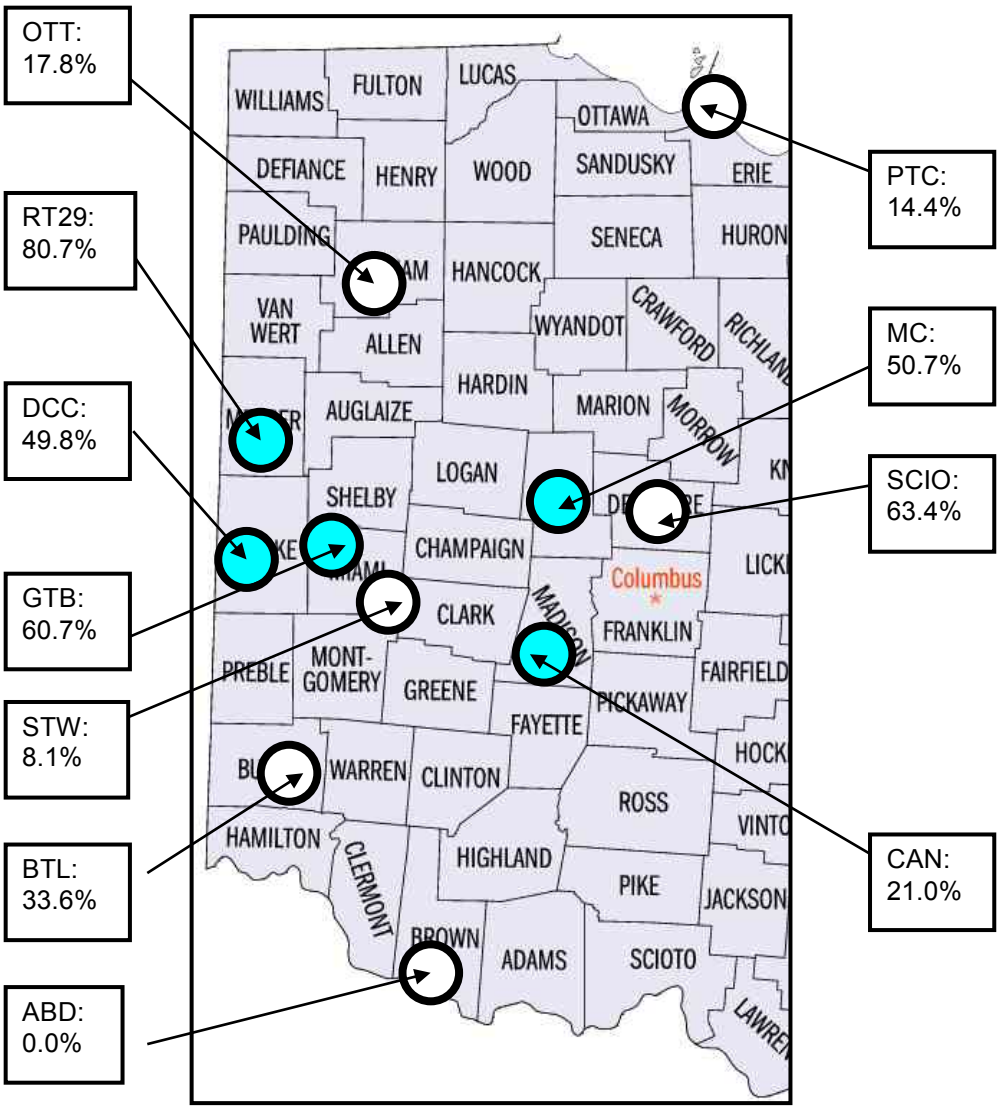

Figure 4.2. Geographical location and herbicide resistance level (from the 1st round of screening) for each Ohio population. Agricultural populations are shown as blue circles, and non-agricultural populations are shown as white circles.

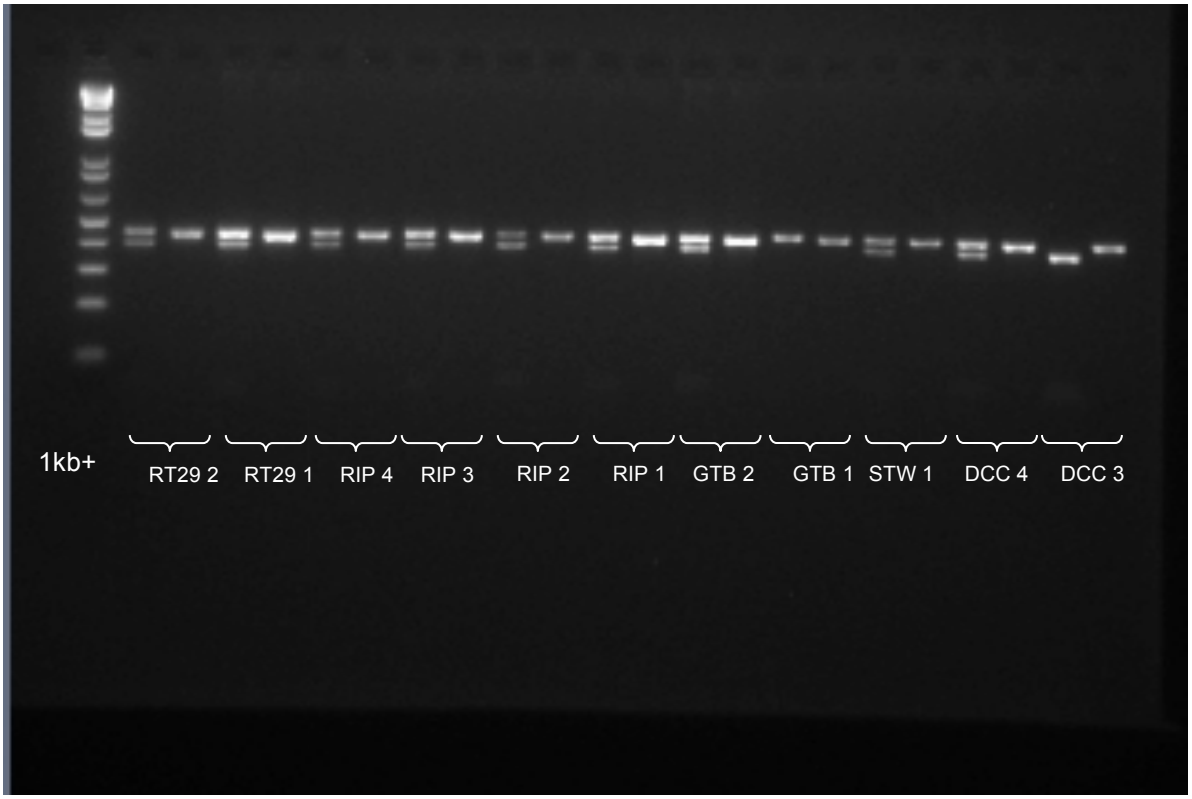

Figure 4.3. Image of 2% agarose gel loaded with 3 uL undigested PCR product (right-hand well for each individual) and 3 uL restriction-enzyme digested PCR product (left-hand well for each individual). Individual samples are labeled, and the 1kb+ ladder is on the left side of the diagram. DCC3 is an example of a homozygote for the Trp574Leu mutation, DCC4 is an example of a heterozygote for the Trp574Leu mutation, and GTB1 is the single sample that lacks the Trp574Leu mutation. RIP refers to an Illinois population not discussed in this chapter.

CONCLUSION OF THE DISSERTATION

This dissertation examined the evolution of a subset of invasive plant species, agricultural weeds, at several different time scales. First, the evolution of agricultural invasiveness on a phylogenetic time scale was studied by testing for associations of morphological and ecological traits with agricultural weediness within a genus. Secondly, the recent invasion of agricultural ecosystems by one species within the genus was investigated with population genetics. Potential intraspecific differences in adaptation of this species to agricultural habitats were tested using a common garden study. Finally, the dynamics of herbicide resistance evolution, an agriculturally-adaptive trait acquired by the weedy intraspecific variety over the last 20 years, was examined by testing for potential herbicide resistance "reservoirs" in natural habitats of the same species. Here I summarize the major findings of each chapter of the dissertation, and their potential impacts on the fields of weed science and invasion biology.

In Chapter 1, I reconstructed the *Amaranthus* phylogeny using six molecular markers to answer questions about the biogeographic relationships and monophyly of the subgenera in the group, as well as to test for phylogenetic signal in *Amaranthus* weed evolution. I found that the monophyly of the three subgenera, Acnida, Albersia, and Amaranthus (as defined in Mosyakin and Robertson, 1996), is not supported. The dioecious species of subgenus Acnida are closely related to a monoecious species, *A. pumilus*, which has never been included in this subgenus. The Eurasian/South African/Australian clade plus its subtending South American species broadly corresponds to subgenus Albersia, except that it does not include the Galápagos species and their close relatives, which were formerly included in Albersia. And the Hybridus Clade includes all of the species usually included in subgenus Amaranthus, but also includes *A. palmeri* and *A. watsonii* (according to the nuclear gene trees), which are dioecious and usually placed into subgenus Acnida. The substantial disagreement between nuclear and chloroplast-based gene

trees in *Amaranthus* is another significant finding of my phylogenetic work: both chloroplast capture in the lineage leading to *A. palmeri/A. watsonii* and incomplete lineage sorting are invoked as explanations for this disagreement.

The biogeographic relationships in *Amaranthus* are also interesting. Although the root of the tree is poorly resolved, the genus appears to have originated in the Americas, and only one major clade in the genus gave rise to Old World species, possibly via a single long-distance dispersal event from South America. Furthermore, the genus colonized the Galápagos Islands in three or four independent events, rather than radiating within the islands. The closest relatives of the Galápagos species are found in western North America or the Caribbean, rather than mainland South America, which fits well with a recent revision of the biogeographic relationships of many other Galápagos plants (Tye and Francisco-Ortega, 2011).

Finally, the tests for trait associations with weediness yielded some expected and some unanticipated results. There is no phylogenetic signal in agricultural invasiveness in *Amaranthus*, which suggests a lack of phylogenetic constraint in the evolution of traits adaptive in agricultural environments. My non-phylogenetic tests showed that agricultural weeds have significantly larger geographic ranges, are more likely to be ruderal (found in waste places), and are more likely to have had their ranges expanded by human activity. Unexpectedly, *Amaranthus* agricultural weed species never occur on beaches (despite the natural disturbance in these habitats), and grow at significantly higher maximum elevations than non-weeds.

In Chapter 2, I used microsatellite markers to test hypotheses about the origin and evolution of the agricultural weed form of *Amaranthus tuberculatus*, or waterhemp, a native Midwestern dioecious species that has invaded agricultural environments within the past 100 years. I found genetic evidence of two ancestral populations within the species, at the western

and eastern ends of the range, validating the observations of Sauer (1957) and Pratt and Clark (2001), although these authors were more concerned with taxonomic issues than with intraspecific variation as an end in itself. My hypothesis that populations of *A. tuberculatus* in agricultural fields were genetically differentiated from populations in natural habitats was not supported, either at a small geographic scale (the St. Louis region), or at larger geographic scales (western Ohio and the entire species range). There is probably too much gene flow between these environments, and too much recombination within the waterhemp genome, for adaptation to these very different habitats to be apparent using neutral markers. Finally, the hypothesis (originally outlined by Sauer in 1957) that the agricultural weed form of *A. tuberculatus* was created by hybridization between the two ancestral taxa (called varieties in this study) was not supported. Instead, it appears that the eastward migration of the western genetic cluster, *A. tuberculatus* var. *rudis*, was the primary factor involved in agricultural invasion in this species.

In Chapter 3, I specifically investigate the hypothesis of varying levels of adaptation to agricultural fields in populations across *Amaranthus tuberculatus'* range, using two soybean common garden plots placed inside and outside of the geographical area where waterhemp is a weed. My prediction was that waterhemp from the most heavily agriculturally-infested "Mississippi Valley" region (MO, IL, and IA) would have higher fitness in agricultural environments than would populations from less infested parts of the species range, and that these plants would also demonstrate local adaptation through their superior performance in the Missouri common garden plot. I found that plants from the Northeastern part of the species range (corresponding to the territory of the ancestral eastern genetic cluster, *A. tuberculatus* var. *tuberculatus*) had unequivocally lower fitness in crop field habitats than did plants from either the Mississippi Valley or Plains (NE, KS, OK) regions. However, the relative performance of

the Mississippi Valley and Plains plants depended on the location of the common garden: in the Missouri plot (within the Mississippi Valley region), Mississippi Valley plants had the highest fitness, whereas in the Ohio plot, average fitness measurements did not differ between plants from the two regions. This finding supports the idea of local adaptation of the Mississippi Valley plants, and also suggests that the western variety, *A. tuberculatus* var. *rudis* (corresponding roughly to the Plains + Mississippi Valley populations) was preadapted to agricultural invasion, rather than requiring genetic and phenotypic changes to be successful in crop fields.

Finally, Chapter 4 asks whether an allele that probably only confers higher fitness in agricultural habitats (ALS-inhibitor resistance) is also found in natural habitats of *Amaranthus tuberculatus*. This question is examined in the western half of Ohio, where only part of the state is infested with agricultural waterhemp (the "agricultural waterhemp region"), and there is the potential to detect long-distance movement of herbicide resistance alleles into populations far from this region. My results show the presence of the same resistance allele (with the Trp574Leu mutation) in agricultural and natural Ohio waterhemp populations, and the correspondence of the frequencies of this allele in populations outside the agricultural waterhemp region with the levels of admixture observed in the same populations in Chapter 2 suggests that gene flow is responsible for this pattern.

Chapter 4 has the greatest implications for the field of weed science, a field that is driven largely by the goal of weed control. Control of waterhemp and another dioecious *Amaranthus* agricultural weed, *A. palmeri*, is already extremely difficult for farmers: currently, there are few chemical classes of herbicides to which at least some populations of *A. tuberculatus* and *A. palmeri* have not evolved resistance (Gaines et al., 2010). The results of my herbicide resistance

research raise the disturbing possibility that at least for mutations lacking a significant fitness cost in the absence of the herbicide (and fitness costs are only conclusively shown for triazine herbicides), herbicide resistance can persist in waterhemp populations outside agricultural ecosystems. *Amaranthus palmeri* also commonly occurs outside crop fields, along railroads, roadsides, and in other anthropogenically-disturbed habitats (Mosyakin and Robertson, 2003), so the same results could apply to this species. These non-agricultural populations could potentially act as genetic reservoirs, or secondary sources, for herbicide resistance alleles, and thus nearby agricultural populations could regain herbicide resistance through gene flow after that particular type of resistance was eradicated from the agricultural habitat. This could prevent older herbicides from ever being effective again to control the weedy dioecious *Amaranthus* species; this is a problematic scenario because farmers are returning to some of the older chemical classes as glyphosate resistance develops in the Midwest and the southern U.S. (Mortensen, 2010).

Weed scientists recommend herbicide rotation and crop rotation, as well as using herbicides in combination, to slow down the evolution of herbicide resistance (Hager and Refsell, 2008). Unfortunately, "superweeds" such as *A. tuberculatus* and *A. palmeri* may force farmers to revert to cultivation for weed control as well, reversing some of the positive improvements in topsoil and nutrient retention that have resulted from conservation tillage. The most important idea for farmers to take from evolutionary studies of agricultural weeds is the idea of varying selection pressure: there is no way to stop weeds from developing resistance to any particular management technique eventually, but adaptation to that technique can certainly be slowed by not imposing enormous selection pressure in a single direction, year after year, on one population of weeds (Neve et al., 2009).

At the genus level, the relatedness of various *Amaranthus* species could affect their

ability to hybridize, and potentially to exchange important "weedy" alleles such as herbicide resistance alleles. However, previous studies on hybridization between *Amaranthus* weeds demonstrate that the level of reproductive compatibility between two species can be hard to predict based on their degree of relationship. *Amaranthus tuberculatus* and *A. palmeri* are as related as *A. tuberculatus* and *A. hybridus* (or more related, at least at chloroplast loci), based on my phylogenetic results. However, stronger pre- and postzygotic barriers between *A. tuberculatus* and *A. palmeri* seem to exist: of the 69 offspring from an experimental cross of these two species, 60 were the result of agamospermy in *A. palmeri*, eight were nonviable, and only one was a true, fertile hybrid (Trucco et al., 2007). On the other hand, *A. tuberculatus* and *A. hybridus* frequently hybridize in nature (Pratt, 1999) and can be successfully crossed in a controlled setting; although fertility is greatly reduced in the hybrids, backcrosses with *A. tuberculatus* can transfer a number of *A. hybridus* alleles into this species (the same is not true for the reciprocal backcross) (Trucco et al., 2009). The phylogeny could be helpful for generating hypotheses about reproductive compatibility between weed species in the genus, but these should be carefully tested with greenhouse experiments.

My findings also have significance for invasion biology. Because of the previous dominance of purely ecological hypotheses in invasion biology, the importance of evolution and especially of hybridization in explaining invasive success has been emphasized in many recent papers (see Lee, 2002; Schierenbeck and Ellstrand, 2009). The results from Chapter 2 of my dissertation show that these processes are not always necessary to produce an invasive species (in accordance with Parker et al., 2003). In some cases, a species (or populations within a species) may be preadapted to invade a new habitat (e.g., Fenesi et al., 2011; Van Kleunen et al., 2011). In these cases, the new habitat may have similarities to the species' natural habitat:

waterhemp evolved to withstand and take advantage of the disturbance dynamics in floodplains in the Midwest, and the natural disturbance regimes in these ecosystems may have important characteristics in common with Midwestern agricultural practices. For instance, germination throughout the growing season is presumably an adaptation that prevents entire populations of waterhemp from being wiped out by a single flood, but it is also very useful in avoiding extermination by a single application of herbicide in crop fields.

Both Chapter 2 and Chapter 3 of the dissertation show that variation within invasive species should be taken into account in evolutionary studies of these taxa. Species are not Platonic ideals, and population structure in genetic variation and adaptive traits means that the success of an introduction can sometimes depend on which population within the native range is the source. This is not an entirely new idea in invasion biology (see Ward et al., 2008), but it is often overlooked even in highly geographically structured species, probably partly because introductions that fail to establish are difficult to document. Chapters 2 and 3 also drive home the importance of changes in the invaded ecosystem itself in allowing new species to compete successfully with the current residents. Natural ecosystems often experience disturbance and climatic fluctuations that might permit invasion at some times and not others (encompassed in the concept of ecosystem invisibility, Lonsdale, 1999), although the significance of any particular change is harder to decipher because of the complexity of these ecosystems relative to agricultural environments.

At the generic level, I found no phylogenetic signal in agricultural weediness in *Amaranthus*. As phylogenetic signal has been found in a variety of genera containing multiple invasive species of natural ecosystems, it is possible that agricultural weeds use a greater variety of mechanisms to invade agricultural ecosystems, even within a genus. Alternatively, genera

that produce agricultural weeds may be less constrained by evolutionary history and thus freer to evolve the same agriculturally invasive traits repeatedly. The fact that I found some trait associations with weediness in *Amaranthus* supports the second hypothesis. As mentioned in Chapter 1, these association tests were conducted mainly to generate interesting hypotheses for further testing. For instance, agricultural weeds in *Amaranthus* may come only from ecosystems with particular types of natural disturbance, given that none are found in beach environments. I know of no studies that examine this idea for invasive species of natural ecosystems (see Lee and Gelembiuk, 2008, for a discussion of disturbance and invasive evolution). If future studies find support for this hypothesis, valuable insight into preadaptation and mechanisms of invasion could be gained.

In terms of significance to the broader field of evolutionary biology, the population level studies in this dissertation suggest that agricultural weeds that are not related to domesticated species, or do not occur sympatrically with related crops, can be just as interesting from an evolutionary perspective as those with close crop relatives. Rapid evolutionary change does not necessarily require genetic input from other species; at least some agricultural weeds are capable of evolving very problematic characteristics (such as herbicide resistance) all on their own. Agricultural practices will continue to change, and in fact two new types of herbicide-resistant crops (corn, soy, and cotton resistant to dicamba and 2,4-D) are currently being reviewed for environmental impacts by the USDA (APHIS statement, 2013). Other changes may be beyond the control of any single farmer, including the likely dramatic impacts of climate change on agricultural weeds (Fuhrer, 2003). New agricultural weeds are undoubtedly waiting in the wings to take advantage of future agricultural revolutions.

LITERATURE CITED

APHIS. 2013. Questions and answers: notice of intent to prepare environmental impact statements for 2,4-D and dicamba genetically engineered plants under consideration for deregulation. United States Department of Agriculture, Animal and Plant Health Inspection Service, Biotechnology Regulatory Services Factsheet, May 2013.

FENESI, A., T. RÉDEI, AND Z. BOTTA-DUKÁT. 2011. Hard traits of three *Bromus* species in their source area explain their current invasive success. *Acta Oecologica* 37: 441-448.

FUHRER, J. 2003. Agroecosystem responses to combinations of elevated CO2, ozone, and global climate change. *Agriculture, Ecosystems and Environment* 97: 1-20.

GAINES, T. A., W. ZHANG, D. WANG, B. BUKUN, S. T. CHISHOLM, D. L. SHANER, S. J. NISSEN, ET AL. 2010. Gene amplification confers glyphosate resistance in *Amaranthus palmeri*. *Proceedings of the National Academy of Sciences* 107: 1029-1034.

HAGER, A. G., AND D. REFSELL. 2008. Chapter 12: Weed resistance to herbicides. *In* S. Bissonnette [ed.], 2008 Illinois Agricultural Pest Management Handbook, University of Illinois Extension, Urbana-Champaign, Illinois, USA.

LEE, C. E. 2002. Evolutionary genetics of invasive species. *Trends in Ecology and Evolution* 17: 386-391.

LEE, C. E., AND G. W. GELEMBIUK. 2008. Evolutionary origins of invasive populations. *Evolutionary Applications* 1: 427-448.

LONSDALE, W. M. 1999. Global patterns of plant invasions and the concept of invisibility. *Ecology* 80: 1522-1536.

MORTENSEN, D. A. 2010. Statement prepared for the Domestic Policy Subcommittee of the Oversight and Government Reform Committee. House of Representatives Hearing, 111 Congress, Second Session: Are superweeds an outgrowth of USDA biotech policy? Available online: http://oversight.house.gov/wpcontent/uploads/2012/01/20100728Mortensen.pdf

MOSYAKIN, S., AND K. R. ROBERTSON. 1996. New infrageneric taxa and combinations in *Amaranthus* (Amaranthaceae). *Annals of Botanici Fennici* 33: 275-281.

MOSYAKIN, S., AND K. R. ROBERTSON. 2003. *Amaranthus*. *In* Flora of North America Editorial Committee [eds.], Flora of North America North of Mexico, Vol. 4. Magnoliophyta: Caryophyllidae, Pt. 1. Oxford University Press, New York, USA.

NEVE, P., M. VILA-AIUB, AND F. ROUX. 2009. Evolutionary-thinking in agricultural weed management. *New Phytologist* 184: 783-793.

PARKER, I. M., J. RODRIGUEZ, AND M. E. LOIK. 2003. An evolutionary approach to understanding the biology of invasions: local adaptation and general-purpose genotypes in the weed *Verbascum thapsus*. *Conservation Biology* 17: 59-72.

PRATT, D. B. 1999. Taxonomic revision of *Amaranthus rudis* and *A. tuberculatus* (Amaranthaceae). Masters Thesis, Iowa State University.

PRATT, D. B, AND L. G. CLARK. 2001. *Amaranthus rudis* and *A. tuberculatus*, one species or two? *Journal of the Torrey Botanical Society* 128: 282-296.

SAUER, J. D. 1957. Recent migration and evolution of the dioecious amaranths. *Evolution* 11: 11-31.

SCHIERENBECK, K. A., AND N. C. ELLSTRAND. 2009. Hybridization and the evolution of invasiveness in plants and other organisms. *Biological Invasions* 11: 1093-1105.

TRUCCO, F., D. ZHENG, A. J. WOODYARD, J. R. WALTER, T. C. TATUM, A. L. RAYBURN, AND P. J. TRANEL. 2007. Nonhybrid progeny from crosses of dioecious amaranths: implications for gene-flow research. *Weed Science* 55: 119-122.

TRUCCO, F., T. TATUM, A. L. RAYBURN, AND P. J. TRANEL. 2009. Out of the swamp: unidirectional hybridization with weedy species may explain the prevalence of *Amaranthus tuberculatus* as a weed. *New Phytologist* 184: 819-827.

TYE, A., AND J. FRANCISCO-ORTEGA. 2011. Origins and evolution of Galápagos endemic vascular plants. *In* D. Bramwell and J. Caujapé-Castells, [eds.], The Biology of Island Floras, 89-153. Cambridge University Press, Cambridge, UK.

VAN KLEUNEN, M., D. R. SCHLAEPFER, M. GLAETTLI, AND M. FISCHER. 2011. Preadapted for invasiveness: do species traits or their plastic response to shading differ between invasive and non-invasive plant species in their native range? *Journal of Biogeography* 38: 1294-1304.

WARD, S. M., S. D. REID, J. HARRINGTON, J. SUTTON, AND K. G. BECK. 2008. Genetic variation in invasive populations of yellow toadflax (*Linaria vulgaris*). *Weed Science* 56: 394-399.